

FEATURES					
source	Location/Qualifiers				
	1..448 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="S3SNUI6-31-B04" /sex="F" /tissue_type="Ascites" /cell_line="Lymphoblast-like" /lab_host="T0P10F" /clone_lib="S3SNUI6"				
	/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoR I which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."				
ORIGIN					
Query Match	19.7%;	Score 343;	DB 12;	Length 448;	
Best Local Similarity	99.7%;	Pred. No. 1.6e-58;			
Matches 393;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
OY	15	AGACCGCTGCGGGCCGCCAGGCGCGGGATGTCCCTGAAATGCGCGGGCAGCGGCG 74			
Db	28	AGACCGCTGCGGGCCGCCAGGCGCGGGATGTCCCTGAAATGCGCGGGCAGCGGCG 87			
OY	75	ACGCGCCCTTGCGAGCCTGTGAACAAGCAACCGCACCCGCTTTCTCTTCGACG 134			
Db	88	ACGCGCCCTTGCGAGCCTGTGAACAAGCAACCGCACCCGCTTTCTCTTCGACG 147			
OY	135	TCAAAGGGGACCAACCGGCTGTGCTGGCGCGGTGAGAACAACGCTGTGGTGCATCT 194			
Db	148	TCAAAGGGGACCAACCGGCTGTGCTGGCGCGGTGAGAACAACGCTGTGGTGCATCT 207			
OY	195	TTGCAGTGTGCTGTGGCAACGTGTGGCCCTGTGTGGTGGCGCGGACGACGCC 254			
Db	208	TTGCAGTGTGCTGTGGCAACGTGTGGCCCTGTGTGGTGGCGCGGACGACGCC 267			
OY	255	GCGGCGGACATGCTGCTGTACTCAACCTTTCTGGCGGACCTGTCTTCATCAGCG 314			
Db	268	GCGGCGGACATGCTGCTGTACTCAACCTTTCTGGCGGACCTGTCTTCATCAGCG 327			
OY	315	CTATCCCTCTGTGCTGTGGCCGTGTGGCTGGAATGAGGCTGTGGCCCCCTTGCT 374			
Db	328	CTATCCCTCTGTGCTGTGGCCGTGTGGCTGGAATGAGGCTGTGGCCCCCTTGCT 387			
OY	375	GCCACCTGCTCTTCAAGCATGACCCCTGAGCGG 408			
Db	388	GCCACCTGCTCTTCAAGCATGACCCCTGAGCGG 421			
RESULT 5					
LOCUS	A1469384/C	442 bp	mRNA	linear	EST 14-APR-1999
DEFINITION	tm08a12.x1 NCI_CGAP_Col4 Homo Sapiens cDNA clone IMAGE:2155966 3'				
KEYWORDS	similar to contains Alu repetitive element;contains element THR repetitive element ; , mRNA sequence.				
ACCSSION	A1469384				
VERSIONS	A1469384.1	GI:4331474			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM					
Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 442) NCI-Gap http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs@remail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmerl-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at: www-bio.llnl.gov/bdrp/image/image.html Insert length: 777 Std Error: 0.00 Seq primer: -40UP from Glibco High quality sequence stop: 401. location/qualifiers				
FEATURES	source				
	1..442 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:2155966" /tissue_type="moderately-differentiated adenocarcinoma" /lab_host="DH108" /clone_lib="NCI CGAP Col4" /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 kb. Life Technologies catalog #: 11531-019"				
ORIGIN					
Query Match	9.9%;	Score 172;	DB 9;	Length 442;	
Best Local Similarity	100.0%;	Pred. No. 6.3e-25;			
Matches 172;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	1286	AAGTTGATAATATTTCCTTTTAAGAAGATTGTGGCCAGTGCAAGTGTTCATGC 1345			
Db	442	AAGTTGATAATATTTCCTTTTAAGAAGATTGTGGCCAGTGCAAGTGTTCATGC 383			
OY	1346	CTGTAAATCCAGCAGATTGGAGGCTGAGGTGGTGATCACTGAGTCAAGATTGCA 1405			
Db	382	CTGTAAATCCAGCAGATTGGAGGCTGAGGTGGTGATCACTGAGTCAAGATTGCA 323			
OY	1406	GACCAACTGACCAACATGTGTGAGACCCCGCTCTCTATAAATAAAAAA 1457			
Db	322	GACCAACTGACCAACATGTGTGAGACCCCGCTCTCTATAAATAAAAAA 271			
RESULT 6					
LOCUS	BF924587	291 bp	mRNA	linear	EST 19-JAN-2001
DEFINITION	IL5-NT0228-271100-271-h06 NT0228 Homo sapiens cDNA, mRNA sequence.				
ACCSSION	BF924587				
VERSION	BF924587.1	GI:12320473			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000).
MEDLINE
20202663
PUBMED
10737800
COMMENT
Contact: Stimpson A.J.G.

Query Match	3.8%	Score 67	DB 10	Length 291
Best Local Similarity	100.0%	Pred. No.	0.00039	
Matches 67	Conservative 0	Mismatches 0	Indels 0	Gaps 0

[illegible]

Email: estewatson.wustl.edu
 Insert Size: 1494
 High quality sequence stops: 370 Source: IMAGE Consortium, LBNL
 This clone is available royalty-free through LBNL ; contact the
 IMAGE Consortium (info@image.lbl.gov) for further information.
 Insert Length: 1494 Std Error: 0.00
 Seq primer: Promega -21m3
 High quality sequence stop: 370.
 Location/Qualifiers

ORIGIN

	Query Match	3.7%	Score 65;	DB 14;	Length 451;
	Best Local Similarity	100.0%	Pred. No. 0.00064;		
	Matches 65;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1344	GCCCTGATATCCACAGCAGTTTGGGAGCGCTGAGTGGTGATCACCCTGAGTCCAGAGATTC			1403
Db	275	GCCCTGATATCCACAGCAGTTTGGGAGCGCTGAGTGGTGATCACCCTGAGTCCAGAGATTC			216
OY	1404	GAGAC 1408			
Db	215	GAGAC 211			

[illegible]

ORIGIN /db xref="taxon:9606"
/clone lib="MAGE resequences, MAGM"
/note="Vector: pBluescriptSKm"

Query Match 3.4%; Score 60; DB 10; Length 602;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1351 ATCCGACAGTTGGAGGCTGAGGTGGATCAGCTGAGCTGAGAGTTGAGACCA 1410
|||||
Db 253 ATCCGACAGTTGGAGGCTGAGGTGGATCAGCTGAGCTGAGAGTTGAGACCA 194

RESULT 9
AA558426 229 bp mRNA linear EST 08-SEP-1997
LOCUS nk38c09.s1 NCI CGAP GC2 Homo sapiens cDNA clone IMAGE:1015792 3'
DEFINITION similar to contain Alu repetitive element; mRNA sequence.
ACCESSION AA558426
VERSION AA558426.1 GI:2328903
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 229)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Mookalk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.

FEATURES
Source
1. .229
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1015792"
/issue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP GC2"
/note="Vector: Bluescript SK-; Site 1: Score; Site 2:
XhoI; Cloned unidirectionally. Primer: 0190 dt. BulK
germ cell tumor. 5' adaptor sequence: 5' GATTTCGACGACG 3'
3' adaptor sequence: 5' CTCGAGTTTTCCTTTTTCCTTTT 3'
Average insert size: 1.2 Kb."

ORIGIN
Query Match 3.4%; Score 59; DB 9; Length 229;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1352 TCCGACAGTTGGAGGCTGAGGTGGATCAGCTGAGCTGAGAGTTGAGACCA 1410
|||||
Db 229 TCCGACAGTTGGAGGCTGAGGTGGATCAGCTGAGCTGAGAGTTGAGACCA 171

RESULT 10
BX506188 446 bp mRNA linear EST 04-SEP-2003
LOCUS BX506188

DEFINITION DKFZp686C04221_r1 686 (synonym: hlc3) Homo sapiens cDNA clone

ACCESSION DKFZp686C04221 5', mRNA sequence.
VERSION BX506188
KEYWORDS BX506188.1 GI:32037167
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 446)
Foustka, A., Albert, R., Moosmayer, P., Schnupp, I., Welleneuther, R.,
Wewes, H.W., Well, B., Aml, C., Oeanger, A., Foto, G., Han, M. and
Wiemann, S.
EST (Poussta, A., Albert, R., Moosmayer, P., Schnupp, I.,
Welleneuther, R., et al.)
Unpublished (2003)
CONTACT: MTPS

TITLE Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
JOURNAL This is the 5' sequence of the clone insert
COMMENT This clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp686C04221) is available at the RZPD in Berlin.
Please contact the RZPD: Reasourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. .446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C04221"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlc3)"
/note="Vector: pTriplex2; Site_1: sf11A; Site_2: sf11B;
cDNA-collection"

ORIGIN

Query Match 3.4%; Score 59; DB 13; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1345 CCTGTATCCGACAGTTGGAGGCTGAGGTGGATCAGCTGAGAGTTTC 1403
|||||
Db 294 CCTGTATCCGACAGTTGGAGGCTGAGGTGGATCAGCTGAGAGTTTC 236

RESULT 11
BE061278 637 bp mRNA linear EST 09-JUN-2000
LOCUS BE061278
DEFINITION I10-BT0168-031199-135-a05 BT0168 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE061278
VERSION BE061278.1 GI:8405928
KEYWORDS EST.
SOURCE Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 637)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663

PUBMED

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer ResearchRua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PAPSP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ct2-ll0-B70168-021

199-135-8056c3-199-11-036c4-1)

Seq primer: puc 18 forward

High quality sequence start: 17
Location/Qualifiers

FEATURES

source

1. 637

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_1lb="B70168"

/note="Organ: breast; Vector: pUC18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN

Query Match 3.3%; Score 57; DB 10; Length 637;

Best Local Similarity 100.0%; Pred. No. 0.018;

Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1340 TCATGCTGTAATCCAGCAGTTGGAGGCTGAGTGGTGCATCCTGAGGTCA 1396

Db 376 TCATGCTGTAATCCAGCAGTTGGAGGCTGAGTGGTGCATCCTGAGGTCA 432

RESULT 12

CD644362/c 779 bp mRNA linear EST 17-JUN-2003

LOCUS CD644362

DEFINITION AGENCOURT_14553420 NIA Human H1 Embryonic Stem Cell cDNA Library

(Long) Homo sapiens cDNA clone IMAGE:30424873 5', mRNA sequence.

ACCESSION CD644362

VERSION CD644362.1 GI:31816168

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 779)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: egsbds-f@mail.nih.gov

Tissue Procurement: Irene Glins and Mahendra Rao, NIA

CDNA Library Preparation: Yulan Piao and Minoru Ko

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC c

One distribution information

can be found through the I.M.A.G.E. Consortium/LMNL at:

http://image.llnl.gov

Plate: NDAM508 row: d column: 02

High quality sequence stop: 650.

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30424873"

/tissue_type="Embryonic Stem cells"

/cell_line="WA01"

/lab_host="DH10B (T1 phage-resistant)"

/clone_1lb="NIA Human H1 Embryonic Stem Cell cDNA Library

(Long)"

/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;

This is a long-transcript enriched cDNA library (Genome

Res. 11: 1553-1558 (2001). [PMID: 11544191] from WA01

cell line. Undifferentiated human ES cell line WA01/H1

was obtained from WICell Research Institute, Inc.,

Madison, WI, cultured according to their instructions, on

MEF feeders. They formed round colonies with defined edges

and were positive for alkaline phosphatase, SSEA-4, OCT3,

OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are

negative for GATP2, GATP4, PDX1, NCM, MSX1, FLT3, SSEA-1,

TUBB3, NES, GFAP, and RORs. When confluent (18-10 days

after plating), the ES cells from 4 x 6cm dishes were

treated with 1 mg/ml collagenase, type IV

(Invitrogen/GIBCO) for 5-10 min and gently scraped off

with 5 ml pipette. RNA was purified with TRIzol Reagent

from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558

(2001). [PMID: 11544191] Double-stranded cDNAs were

synthesized with an Oligo(dT) primer (Invitrogen:

5'-pGACTAGTCTTAATCCAGCAGTTGGAGGCTGAGTGGTGCATCCTGAGGTCA

3'-) from total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated

to linc-linker IV-SalI, purified by phenol/chloroform

extraction, and separated from free linkers by

Centricon-100 column. Then, the cDNAs were amplified by

long-range high fidelity PCR using Ex Taq polymerase

(Takara) with a primer SalI-S for 25 cycles. The products

were purified by phenol/chloroform extraction and

Centricon-100 column. The cDNAs were digested with SalI

and NotI enzymes and cloned into SalI/NotI site of

pCMV-Sport6 plasmid vector. The average insert size is

about 3.6kb."

ORIGIN

Query Match 3.3%; Score 57; DB 14; Length 779;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1340 TCATGCTGTAATCCAGCAGTTGGAGGCTGAGTGGTGCATCCTGAGGTCA 1396

Db 356 TCATGCTGTAATCCAGCAGTTGGAGGCTGAGTGGTGCATCCTGAGGTCA 300

RESULT 13

CB448183 637 bp mRNA linear EST 26-MAR-2003

LOCUS CB448183

DEFINITION 7022244 MARRC 6BOV Bos taurus cDNA.5', mRNA sequence.

ACCESSION CB448183

VERSION CB448183.1 GI:2934565

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 637)

Smith,T.P.L., Roberts,A.J., Echterkamp,S.E., Chitko-McKown,C.G.,

Wray,J.B. and Keeler,J.W.

A second set of bovine ESTs from pooled-tissue normalized libraries

Unpublished (2003)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

FEATURES

source

1. 779

Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.

Plate: F0Y8057 row: L column: 17
Seq primer: GTAATACGACTCACTATAGG3.

Location/Qualifiers

1. 637

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN

Query Match 3.2%; Score 56; DB 14; Length 637;
Best Local Similarity 100.0%; Pred. No. 0.028; Indels 0; Gaps 0;
Matches 56; Conservative 0; Mismatches 0;

Qy 857 TCCTTCTTCATCATGTGAGCCCATCATCATCATCTCTCTCATCTGATCCA 912
|||||
260 TCCTTCTTCATCATGTGAGCCCATCATCATCATCTCTCTCATCTGATCCA 315

RESULT 14
BC017579 1287 bp mRNA linear HTC 04-MAR-2003
LOCUS Homo sapiens, clone IMAGE:4107433, mRNA.

DEFINITION BC017579
ACCESSION BC017579
VERSION BC017579.1 GI:17068402
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1287)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
<http://www.systemsbioology.org>
contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whitting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 32 Row: C Column: 11
This clone has the following problem: retained intron.

location/Qualifiers

1. 1287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4107433"
/tissue_type="bone marrow, acute myelogenous leukemia"
/clone_lib="NIH_MGC_55"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

ORIGIN

Query Match 3.2%; Score 56; DB 11; Length 1287;
Best Local Similarity 100.0%; Pred. No. 0.015; Indels 0; Gaps 0;
Matches 56; Conservative 0; Mismatches 0;

Qy 1516 CAGAGAAATCTCTTGAACCTGGAGGACGAGTTGACATGAGCCGAGATGTGCCA 1571
|||||
Db 1164 CAGAGAAATCTCTTGAACCTGGAGGACGAGTTGACATGAGCCGAGATGTGCCA 1219

RESULT 15
BU660876 208 bp mRNA linear EST 30-SEP-2002
LOCUS Erythroid Precursor Cells (LCB:c1 library) Homo
DEFINITION c165d10.21 Hembase; Erythroid Precursor Cells (LCB:c1 library) Homo
sapiens cDNA clone c165d10 5', mRNA sequence.

ACCESSION BU660876
VERSION BU660876.1 GI:23373058
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 208)
AUTHORS Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.
TITLE Gene Expression in Human Erythroid Precursor Cells
JOURNAL Unpublished (2002)
COMMENT Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jm7ef@nih.gov

The c1 library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
<http://hembase.nidk.nih.gov>
Plate: 65 row: d column: 10
Seq primer: 5' lambda-Triplex2 Sequencing Primer.
Location/Qualifiers

FEATURES
source

1. 208
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c165d10"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPR++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:c1
library)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI. A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA library
Construction Kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/l peptide nucleic acid
(PNA) oligos
(N-terminal) -biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and
(N-terminal) -biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with SfiI and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC).
[Http://www.nisc.nih.gov/](http://www.nisc.nih.gov/)."

ORIGIN

Query Match 3.2%; Score 55; DB 13; Length 208;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 CTTGAACCTGGAGGCAGAGGTTGCAGTGCAGCCGAGATCCGTCATTCGCACTCCA 1581
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 128 CTTGAACCTGGAGGCAGAGGTTGCAGTGCAGCCGAGATCCGTCATTCGCACTCCA 182

Search completed: October 1, 2004, 03:51:44
 Job time : 4109.41 secs

Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: September 30, 2004, 21:45:48 ; Search time 123.742 Seconds
(without alignments)
6996.194 Million cell updates/sec

Title: US-10-077-698-5

Sequence: 1 ttcgcaagctcagcgaagc.....aaaaaaaaaaaaaaaa 1560

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCOMB.COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1560	100.0	1560	US-09-261-5998-5	Sequence 5, Appl1
2	1560	100.0	1560	US-09-456-455A-5	Sequence 5, Appl1
3	181	11.6	181	US-09-456-455A-14	Sequence 15, Appl1
4	138	8.8	138	US-09-456-455A-15	Sequence 15, Appl1
5	44	2.8	1743	US-09-261-5998-2	Sequence 2, Appl1
6	44	2.8	1743	US-09-456-455A-2	Sequence 2, Appl1
7	38	2.4	241	US-09-328-111-472	Sequence 472, App
8	26	1.7	753	US-09-232-191-14	Sequence 11, Appl1
9	26	1.7	753	US-09-232-200-14	Sequence 11, Appl1
10	26	1.7	753	US-09-232-197-14	Sequence 11, Appl1
11	26	1.7	753	US-09-232-201-14	Sequence 11, Appl1
12	26	1.7	753	US-09-232-195-14	Sequence 11, Appl1
13	26	1.7	1560	US-09-261-5998-5	Sequence 5, Appl1
14	26	1.7	1560	US-09-456-455A-5	Sequence 5, Appl1
15	26	1.7	16442	US-08-781-891-208	Sequence 208, App
16	26	1.7	16442	US-09-618-166-208	Sequence 208, App
17	25	1.6	624	US-09-385-982-414	Sequence 414, App
18	25	1.6	624	US-09-148-545-23	Sequence 23, Appl1
19	25	1.6	960	US-09-248-335-57	Sequence 57, Appl1
20	25	1.6	994	US-09-366-887A-5	Sequence 5, Appl1
21	25	1.6	1174	US-10-154-595-49	Sequence 49, Appl1
22	25	1.6	1243	US-09-731-924A-1	Sequence 1, Appl1
23	25	1.6	1244	US-08-916-442-1	Sequence 1, Appl1
24	25	1.6	1244	US-09-317-641-1	Sequence 1, Appl1
25	25	1.6	1513	US-09-716-129-23	Sequence 23, Appl1
26	25	1.6	1545	US-09-559-023-1	Sequence 1, Appl1
27	25	1.6	1594	US-09-016-434-1059	Sequence 1059, Ap

28	25	1.6	1799	3	US-09-329-633A-1	Sequence 1, Appl1
29	25	1.6	1799	4	US-09-079-029-2	Sequence 2, Appl1
30	25	1.6	2140	3	US-09-058-389A-6	Sequence 6, Appl1
31	25	1.6	2140	4	US-09-611-781-6	Sequence 6, Appl1
32	25	1.6	2880	3	US-09-115-954-3	Sequence 3, Appl1
33	25	1.6	2920	4	US-10-158-847-137	Sequence 137, App
34	25	1.6	3842	3	US-09-115-954-7	Sequence 7, Appl1
35	25	1.6	3912	3	US-09-115-954-1	Sequence 1, Appl1
36	25	1.6	11050	4	US-10-204-708-86	Sequence 86, Appl1
37	25	1.6	49136	3	US-09-422-869-1	Sequence 1, Appl1
38	25	1.6	81001	4	US-09-750-580-1	Sequence 1, Appl1
39	25	1.6	148567	4	US-09-801-876B-3	Sequence 3, Appl1
40	25	1.6	148567	4	US-10-254-869-3	Sequence 3, Appl1
41	25	1.6	152331	3	US-09-128-155-16	Sequence 15, Appl1
42	25	1.6	176373	3	US-09-128-155-17	Sequence 17, Appl1
43	25	1.6	246240	2	US-08-724-394A-20	Sequence 20, Appl1
44	25	1.6	246240	2	US-08-724-394A-21	Sequence 21, Appl1
45	25	1.6	246240	2	US-08-724-394A-22	Sequence 22, Appl1

ALIGNMENTS

RESULT 1
US-09-261-5998-5
Sequence 5, Application US/09261599B

Patent No. 6395877

GENERAL INFORMATION:

APPLICANT: Gluckmann, Maria A.

TITLE OF INVENTION: 14273 Receptor, A No. 6395877 G-Protein Coupled Receptor

FILE REFERENCE: 5800-48, 035800/177086

CURRENT APPLICATION NUMBER: US/09/261,599B

CURRENT FILING DATE: 1999-02-26

PRIOR FILING DATE: 1998-06-30

PRIOR APPLICATION NUMBER: 09/107,761

PRIOR FILING DATE: 1998-12-30

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 5

LENGTH: 1560

TYPE: DNA

ORGANISM: Murine ortholog

US-09-261-599B-5

Query Match 100.0%; Score 1560; DB 4; Length 1560;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTGCCAAGCTCAGCGTGAAGCTCTTCCACTGCATCTCACAGAGGGGTTCAATGAGTGC	60
Db	1	TTGCCAAGCTCAGCGTGAAGCTCTTCCACTGCATCTCACAGAGGGGTTCAATGAGTGC	60
Qy	61	TTTACACCATGATGACCATCTGAGCTTCCGGCTTACCGGAATCTTCACAGGGAG	120
Db	61	TTTACACCATGATGACCATCTGAGCTTCCGGCTTACCGGAATCTTCACAGGGAG	120
Qy	121	TCGATGACCTCTTGAAGCAGCAGAGCGCGAGCTCCGCATCTTCCGAGCGCTGG	180
Db	121	TCGATGACCTCTTGAAGCAGCAGAGCGCGAGCTCCGCATCTTCCGAGCGCTGG	180
Qy	181	GCCGGGCGCCGCGATGTCCTCTGAGTGTGACAGAGAGCGGCGCTGCTCCGACA	240
Db	181	GCCGGGCGCCGCGATGTCCTCTGAGTGTGACAGAGAGCGGCGCTGCTCCGACA	240
Qy	241	CCCTGACCAAGTCATCGACCACTCCCTTCTTCTGAGTGTGACAGAGCGGAGCACC	300
Db	241	CCCTGACCAAGTCATCGACCACTCCCTTCTTCTGAGTGTGACAGAGCGGAGCACC	300
Qy	301	GGTTGGTGTGAGCGTGTGAGAGCAGCCGTTGTGAGACTCATCTTGTGTCTCACTGC	360
Db	301	GGTTGGTGTGAGCGTGTGAGAGCAGCCGTTGTGAGACTCATCTTGTGTCTCACTGC	360

QY 361 TGGGCAACGTGTGTCTCTAGTGTGGTGCGCGCCGTGCGGCGCGTGGGGCGTCAAGCA 420
DB 361 TGGGCAACGTGTGTCTCTAGTGTGGTGCGCGCCGTGCGGCGCGTGGGGCGTCAAGCA 420
QY 421 GCGTGTGTCTCAACCTCTTCTGCGGGATTTGCTTTTCAACAGCGCAATCCCTCTAGTGC 480
DB 421 GCGTGTGTCTCAACCTCTTCTGCGGGATTTGCTTTTCAACAGCGCAATCCCTCTAGTGC 480
QY 481 TCGTGTGTGTCTGATCTGAGGCGCTGCTGTTGGGGGCGCGTGTCTGCAACCTGCTCTTCT 540
DB 481 TCGTGTGTGTCTGATCTGAGGCGCTGCTGTTGGGGGCGCGTGTCTGCAACCTGCTCTTCT 540
QY 541 ACGTGATGACATATGAGGCGGCGCTCAAGATCTCACTCACTGCGCGCGGTGACCTTGAGC 600
DB 541 ACGTGATGACATATGAGGCGGCGCTCAAGATCTCACTCACTGCGCGCGGTGACCTTGAGC 600
QY 601 GCATGTGTGTGTCTGCGGCTCCGCGCGGCGCTTGAAGCGGCGCGGCGGCGGCACTAGG 660
DB 601 GCATGTGTGTGTCTGCGGCTCCGCGCGGCGCTTGAAGCGGCGGCGGCGGCGGCACTAGG 660
QY 661 CCGGACTGTGTGTCTGATATGAGGCTTACCTGCGGCGCTGCGCGGCGCTGCGCTCTCAACT 720
DB 661 CCGGACTGTGTGTCTGATATGAGGCTTACCTGCGGCGCTGCGCGGCGCTGCGCTCTCAACT 720
QY 721 TGTTCGCGGTGTCTCCGCGCGGCGCTTCCGCGGCGGCGGCAAGAAATTCGATTTGCACT 780
DB 721 TGTTCGCGGTGTCTCCGCGCGGCGCTTCCGCGGCGGCGGCAAGAAATTCGATTTGCACT 780
QY 781 TGGATTTGGCCCAACCGCATAGAGAAATCTCATGGATGTGTTTTTGAAGCTTTGAAC 840
DB 781 TGGATTTGGCCCAACCGCATAGAGAAATCTCATGGATGTGTTTTTGAAGCTTTGAAC 840
QY 841 TCCGCGTGTGTGTCTGATATGAGGCTTACCTGCGGCGCTGCGCGGCGCTGCGCTCTCAACT 900
DB 841 TCCGCGTGTGTGTCTGATATGAGGCTTACCTGCGGCGCTGCGCGGCGCTGCGCTCTCAACT 900
QY 901 CATGCGGAAAGAGCTTACGCTGAGCTGAGCTACTCTGAGGCGCAACGATTCGAGTGT 960
DB 901 CATGCGGAAAGAGCTTACGCTGAGCTTGGCATCTGAGAGCGCAACGATTCGAGTGT 960
QY 961 CCCAACAAGACTACCGACTCTTCCGCGCGCTCTTCTGCTCATGTTCTTCTTCACTCA 1020
DB 961 CCCAACAAGACTACCGACTCTTCCGCGCGCTCTTCTGCTCATGTTCTTCTTCACTCA 1020
QY 1021 TGTGAGTGTGTGTCTGATATGAGGCTTACCTGCGGCGCTTGTGATCAAACTTCCGCGAGAGC 1080
DB 1021 TGTGAGTGTGTGTCTGATATGAGGCTTACCTGCGGCGCTTGTGATCAAACTTCCGCGAGAGC 1080
QY 1081 TGTGATCTGTGCTATCCCTTTCTTCTGAGTGTGAGCTTCAAGTTCGCAACTCTGCGC 1140
DB 1081 TGTGATCTGTGCTATCCCTTTCTTCTGAGTGTGAGCTTCAAGTTCGCAACTCTGCGC 1140
QY 1141 TAAACCCCATCTGTACAACTGTGCTGTCAAGAAAGATGAGAAATTTTGTGCT 1200
DB 1141 TAAACCCCATCTGTACAACTGTGCTGTCAAGAAAGATGAGAAATTTTGTGCT 1200
QY 1201 GCTTCTTTTTCGAGAGGAGGCAATTTTACAGATAGCTGTGTCAGGAGGAATGACT 1260
DB 1201 GCTTCTTTTTCGAGAGGAGGCAATTTTACAGATAGCTGTGTCAGGAGGAATGACT 1260
QY 1261 TGTCTGTATTTCCAGCTTAACTAGCTCTGTGTCAGGAGGAACAGGCTGTGCAATGTA 1320
DB 1261 TGTCTGTATTTCCAGCTTAACTAGCTCTGTGTCAGGAGGAACAGGCTGTGCAATGTA 1320
QY 1321 GGGAGTTAATTCAAGAGAAAGCCCAAGTGCCTGCTTTAAATAACCGACTTCCA 1380
DB 1321 GGGAGTTAATTCAAGAGAAAGCCCAAGTGCCTGCTTTAAATAACCGACTTCCA 1380
QY 1381 ACAGCAGGCACTGAGGAGGAGCAAAATTAAGAAATGATGCTGATTAATAATTTT 1440
DB 1381 ACAGCAGGCACTGAGGAGGAGCAAAATTAAGAAATGATGCTGATTAATAATTTT 1440
QY 1441 TCTTAAAGAACTTCTATGGGTCTTTTGTGAACCTTTTAAAGTGTGTAATAT 1500

DB 1441 TCTTAAAGAACTTCTATGGGTCTTTTGTGAACCTTTTAAAGTGTGTAATAT 1500
QY 1501 GATCTAGTAAATTAATTTTATTAATTAAGCTGTTCCTACAAAAAATTTTAAAAA 1560
DB 1501 GATCTAGTAAATTAATTTTATTAATTAAGCTGTTCCTACAAAAAATTTTAAAAA 1560

RESULT 2
US-09-456-455A-5
; Sequence 5, Application US/09456455A
; Patent No. 6448005
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OR INVENTION: 14273 Receptor, A No. 6448005el G-Protein Coupled Receptor
; FILE REFERENCE: NMI-204CP3
; CURRENT APPLICATION NUMBER: US/09/456,455A
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: 09/107,761
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/223,538
; PRIOR FILING DATE: 1998-12-30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Murine ortholog
US-09-456-455A-5

Query Match 100.0%; Score 1560; DB 4; Length 1560;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTGCAAGCTCAAGCTAAGCTCTTCTCACTGCAATCTCAAGAAAGGAGTTTCAGAGTGC 60
DB 1 TTTGCAAGCTCAAGCTAAGCTCTTCTCACTGCAATCTCAAGAAAGGAGTTTCAGAGTGC 60
QY 61 TTTCACACATCAAGAGACACATCCAGACTTGTCCGCGCTTTACCGGAATCTTCAAGCGGAG 120
DB 61 TTTCACACATCAAGAGACACATCCAGACTTGTCCGCGCTTTACCGGAATCTTCAAGCGGAG 120
QY 121 TCGATGACCTCTTTGACAGCCAGAGCGCGGAGCTCCGCACTTTCCGCGAGCGCTG 180
DB 121 TCGATGACCTCTTTGACAGCCAGAGCGCGGAGCTCCGCACTTTCCGCGAGCGCTG 180
QY 181 GCCGGGCGCCCGGAGTGTCTGAGTGTGCAAGAGCGAGCGCTGTGCTTCGCA 240
DB 181 GCCGGGCGCCCGGAGTGTCTGAGTGTGCAAGAGCGAGCGCTGTGCTTCGCA 240
QY 241 CCTGAGCAAGTCAATCGACCCACTTCCCTTTCTCTGGAGTCAAGGCGAGCAACC 300
DB 241 CCTGAGCAAGTCAATCGACCCACTTCCCTTTCTCTGGAGTCAAGGCGAGCAACC 300
QY 301 GGTGTGTGTGAGCGTGTGAGAGCAACGCTTCTGAGACTCATTTGTGTCTCACTGC 360
DB 301 GGTGTGTGTGAGCGTGTGAGAGCAACGCTTCTGAGACTCATTTGTGTCTCACTGC 360
QY 361 TGGGCAACGTGTGTCTCTAGTGTGGTGCGCGCTGCGGCGCGTGGGGCGTCAAGCA 420
DB 361 TGGGCAACGTGTGTCTCTAGTGTGGTGCGCGCTGCGGCGCGTGGGGCGTCAAGCA 420
QY 421 GCGTGTGTCTCAACCTCTTCTGCGGGATTTGCTTTTCAACAGCGCAATCCCTCTAGTGC 480
DB 421 GCGTGTGTCTCAACCTCTTCTGCGGGATTTGCTTTTCAACAGCGCAATCCCTCTAGTGC 480
QY 481 TCGTGTGTGTCTGATCTGAGGCGCTGCTGTTGGGGCGCGTGTGCAACCTGCTCTTCT 540
DB 481 TCGTGTGTGTCTGATCTGAGGCGCTGCTGTTGGGGCGCGTGTGCAACCTGCTCTTCT 540
QY 541 ACGTGATGACATATGAGGCGGCGCTCAAGATCTCACTGAGCGCGGCTGAGAGC 600

Qy 1029 CCCATCATCATCACCATC 1046
Db 121 CCCATCATCATCACCATC 138

RESULT 5
US-09-261-599B-2
Sequence 2, Application US/09261599B
Patent No. 6395877
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 14273 Receptor, A No. 6395877 G-Protein Coupled Receptor
FILE REFERENCE: 5800-4B, 035800/177086
CURRENT APPLICATION NUMBER: US/09/261,599B
CURRENT FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 09/107,761
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: 09/223,538
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1743
TYPE: DNA
ORGANISM: Homo sapiens
US-09-261-599B-2

Query Match 2.8%; Score 44; DB 4; Length 1743;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 852 GGACTGGTCATTGGATCAGTACTCCAAATTTTACAGATCAC 895
Db 701 GGACTGGTCATTGGATCAGTACTCCAAATTTTACAGATCAC 744

RESULT 6
US-09-456-455A-2
Sequence 2, Application US/09456455A
Patent No. 6448005
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 14273 Receptor, A No. 6448005 G-Protein Coupled Receptor
FILE REFERENCE: NMI-204CP3
CURRENT APPLICATION NUMBER: US/09/456,455A
CURRENT FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: 09/107,761
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: 09/223,538
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1743
TYPE: DNA
ORGANISM: Homo sapiens
US-09-456-455A-2

Query Match 2.8%; Score 44; DB 4; Length 1743;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 852 GGACTGGTCATTGGATCAGTACTCCAAATTTTACAGATCAC 895
Db 701 GGACTGGTCATTGGATCAGTACTCCAAATTTTACAGATCAC 744

RESULT 7
US-09-328-111-472
Sequence 472, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:

APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Aetle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 472
LENGTH: 241
TYPE: DNA
ORGANISM: Homo sapiens
US-09-328-111-472

Query Match 2.4%; Score 38; DB 3; Length 241;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 ATCTCAGAGAGGGTTCATGAGTCTTCACACCATC 71
Db 101 ATCTCAGAGAGGGTTCATGAGTCTTCACACCATC 138

RESULT 8
US-09-232-191-14
Sequence 14, Application US/09232191
Patent No. 6284487
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
TITLE OF INVENTION: Fatty Acid Transport Proteins
FILE REFERENCE: WHI97-21D3ME
CURRENT APPLICATION NUMBER: US/09/232,191
CURRENT FILING DATE: 1999-01-14
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-07-20
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 753
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-191-14

Query Match 1.7%; Score 26; DB 3; Length 753;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1535 CCTACAAAAA 1560
Db 726 CCTACAAAAA 751

RESULT 9
US-09-232-200-14

Sequence 14, Application US/09232200A
Patent No. 6288213
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
FILE REFERENCE: WHI97-21P3MB
CURRENT APPLICATION NUMBER: US/09/232,200A
CURRENT FILING DATE: 1999-01-14
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-07-20
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 753
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-200-14

Query Match 1.7%; Score 26; DB 3; Length 753;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1535 CCTACCAAAAAAAAAAAAAAAAAA 1560
DB 726 CCTACCAAAAAAAAAAAAAAAAAA 751

RESULT 10
US-09-232-197-14
Sequence 14, Application US/09232197A
Patent No. 6300096
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
FILE REFERENCE: WHI97-21P3MA
CURRENT APPLICATION NUMBER: US/09/232,197A
CURRENT FILING DATE: 1999-01-14
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-07-20
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 753
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-197-14

Query Match 1.7%; Score 26; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1535 CCTACCAAAAAAAAAAAAAAAAAA 1560
DB 726 CCTACCAAAAAAAAAAAAAAAAAA 751

RESULT 11

US-09-232-201-14
Sequence 14, Application US/09232201A
Patent No. 6348321
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
FILE REFERENCE: WHI97-21P3MC
CURRENT APPLICATION NUMBER: US/09/232,201A
CURRENT FILING DATE: 1999-01-14
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-07-20
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 753
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-201-14

Query Match 1.7%; Score 26; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1535 CCTACCAAAAAAAAAAAAAAAAAA 1560
DB 726 CCTACCAAAAAAAAAAAAAAAAAA 751

RESULT 12
US-09-232-195-14
Sequence 14, Application US/09232195A
Patent No. 6657049
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
FILE REFERENCE: WHI97-21P3MD
CURRENT APPLICATION NUMBER: US/09/232,195A
CURRENT FILING DATE: 1999-01-04
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-07-20
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 753
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-195-14

Query Match 1.7%; Score 26; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1535 CCTACCAAAAAAAAAAAAAAAAAA 1560
DB 726 CCTACCAAAAAAAAAAAAAAAAAA 751

```
RESULT 13
US-09-261-5998-5/c
; Sequence 5, Application US/092615998
; Patent No. 6395877
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14273 Receptor, A No. 639587761 G-Protein Coupled Receptor
; FILE REFERENCE: 5800-4B, 035800/177086
; CURRENT APPLICATION NUMBER: US/09/261,599B
; CURRENT FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/107,761
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/223,538
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Murine ortholog
US-09-261-5998-5

Query Match
Best Local Similarity 1.7%; Score 26; DB 4; Length 1560;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGCCAGCTCAGCGTACGCTCTTCC 27
Db 932 TGCCAGCTCAGCGTACGCTCTTCC 907

RESULT 14
US-09-456-455A-5/c
; Sequence 5, Application US/09456455A
; Patent No. 6448005
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14273 Receptor, A No. 644800561 G-Protein Coupled Receptor
; FILE REFERENCE: NMI-204CP3
; CURRENT APPLICATION NUMBER: US/09/456,455A
; CURRENT FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: 09/107,761
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/223,538
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Murine ortholog
US-09-456-455A-5

Query Match
Best Local Similarity 1.7%; Score 26; DB 4; Length 1560;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGCCAGCTCAGCGTACGCTCTTCC 27
Db 932 TGCCAGCTCAGCGTACGCTCTTCC 907

RESULT 15
US-08-781-891-208
; Sequence 208, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-Bu
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
```

```
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 208:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-781-891-208

Query Match
Best Local Similarity 1.7%; Score 26; DB 3; Length 1642;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1535 CCTACAAAAAAAAAAAAAAAAAAAAA 1560
Db 11978 CCTACAAAAAAAAAAAAAAAAAAAAA 12003

Search completed: October 1, 2004, 03:56:18
Job time : 126.742 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 19:57:17 ; Search time 3672.59 Seconds
(without alignments)
12684.514 Million cell updates/sec

Title: US-10-077-698-5
Perfect score: 1560
Sequence: 1 tgcgcagctcgcgtcagc.....aaaaaaaaaaaaaaaaaaaa 1560

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_esthum:*
- 16: em_estom:*
- 17: em_gse_hum:*
- 18: em_gse_iny:*
- 19: em_gse_pln:*
- 20: em_gse_vtc:*
- 21: em_gse_fun:*
- 22: em_gse_mam:*
- 23: em_gse_mus:*
- 24: em_gse_pro:*
- 25: em_gse_rtd:*
- 26: em_gse_png:*
- 27: em_gse_vtl:*
- 28: gb_gse1:*
- 29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	602	38.6	692	9	AV021454 AV021454
2	557	35.7	721	9	AV025152 AV025152
3	525	33.7	927	13	BO939003 AGENCOURT
4	481	30.8	870	13	BU522456 AGENCOURT

5	465	29.8	875	13	BU522608
6	445	28.5	938	13	BO936343
7	434	27.8	537	10	BX527210
8	420	26.9	671	13	BB609925
9	408	26.2	517	29	CG557110
10	382	24.5	637	10	BB609814
11	278	17.8	437	9	A1552415
12	269	17.2	432	29	CG617050
13	214	13.7	647	10	BF580577
14	194	12.4	345	9	AA030752
15	180	11.5	354	13	BY345055
16	157	10.1	377	9	AV017689
17	154	9.9	318	13	BY346155
18	154	9.9	439	9	AA647793
19	151	9.7	315	13	BY346508
20	129	8.3	376	13	BY344852
21	112	7.2	653	13	BQ256378
22	112	7.2	826	14	CF587177
23	109	7.0	265	9	AV015214
24	109	7.0	650	13	BQ256373
25	102	6.5	228	9	AV291417
26	102	6.5	519	10	BF454386
27	101	6.5	302	13	BY068200
28	101	6.5	309	13	BY052181
29	101	6.5	313	13	BY113659
30	101	6.5	334	13	BY119108
31	101	6.5	336	13	BY145350
32	101	6.5	340	13	BY113754
33	101	6.5	341	13	BY117611
34	101	6.5	341	13	BY124151
35	101	6.5	344	13	BY048577
36	101	6.5	347	13	BY138060
37	101	6.5	350	13	BY098455
38	101	6.5	356	13	BY084277
39	101	6.5	359	13	BY097877
40	101	6.5	359	13	BY085055
41	101	6.5	361	13	BY790047
42	101	6.5	362	13	BY106909
43	101	6.5	366	13	BY077564
44	101	6.5	372	13	BY037967
45	101	6.5	372	13	BY131548

ALIGNMENTS

RESULT 1
AV021454 692 bp mRNA linear EST 23-OCT-2001
LOCUS AV021454 Mus musculus 18-day embryo C57BL/6J Mus musculus CDNA
DEFINITION AV021454.2 GI:16356377
ACCESSION AV021454
VERSION AV021454.2
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Komori, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.
Unpublished (2001)
Unpublished (2001)
On May 11, 1999 this sequence version replaced gi:4798446.
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

TITLE
JOURNAL
COMMENT

Tel: 81-298-36-9145
 Fax: 81-298-36-9098
 Email: genome-res@rken.go.jp
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A.
 and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamana, K.,
 Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
 Hayashizaki, Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences *Mamm. Genome*, 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES
 source
 Location/Qualifiers
 1..692
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1190023803"
 /sex="mixed"
 /dev_stage="18-day embryo"
 /clone_id="Mus musculus 18-day embryo C57BL/6J"

ORIGIN
 Query Match 38.6%; Score 602; DB 9; Length 692;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 842 CCTGTGCGCGGACTGGTCATTTGATCAGTTACTCCAAATTTTACAGATCAGAAAGC 901
 |||||
 89 CTTGTGCGCGGACTGGTCATTTGATCAGTTACTCCAAATTTTACAGATCAGAAAGC 148
 |||||
 902 ATGCGGAAAGGCTTACGCTGAGCTTGGCATCTTGAAGCCACAGATCCGAGTGC 961
 |||||
 149 ATGCGGAAAGGCTTACGCTGAGCTTGGCATCTTGAAGCCACAGATCCGAGTGC 208
 |||||
 962 CCAAGAAGTACGAGCTTTCCGCAAGCTTCTCTGCTCATAGTGTCTTCTTCATCAT 1021
 |||||
 209 CCAAGAAGTACGAGCTTTCCGCAAGCTTCTCTGCTCATAGTGTCTTCTTCATCAT 268
 |||||
 1022 GTGAGTCCATCATCATCAACATCTCTCATCTTGATCCAAATTTCCGGCAGAGACT 1081
 |||||
 269 GTGAGTCCATCATCATCAACATCTCTCATCTTGATCCAAATTTCCGGCAGAGACT 328
 |||||
 1082 GGTTCATTTGGCATCCTTTTCTTGGGTGGTGGCTTCAAGTTGGCAAATCTGCGCCT 1141
 |||||
 329 GGTTCATTTGGCATCCTTTTCTTGGGTGGTGGCTTCAAGTTGGCAAATCTGCGCCT 388
 |||||
 1142 AAAGCCCATCTGTAACAACATGTCGTTGAGGAGAAAGTAAGGAAATTTTCTG 1201
 |||||
 389 AAAGCCCATCTGTAACAACATGTCGTTGAGGAGAAATGAGGAAATTTTCTG 448
 |||||
 1202 CTTCCTTTTCCAGAGAGGAGCCATTTTACAGATACGTCGTGAGGCGAAATGACTT 1261

|||||
 Db 449 CTTCCTTTTCCAGAGAGGAGCCATTTTACAGATACGTCGTGAGGCGAAATGACTT 508
 |||||
 Oy 1262 GTCTGTTATTTCCAGCTTACTAGCTTGTGTCAGGTGAACACCGTGTGCATGTTAAG 1321
 |||||
 Db 509 GTCTGTTATTTCCAGCTTACTAGCTTGTGTCAGGTGAACACCGTGTGCATGTTAAG 568
 |||||
 Oy 1322 GGAATTACTTCAAGGAAGCCCAACAGTCGCGCTCTTAAATAATACCGACTTCCAA 1381
 |||||
 Db 569 GGAATTACTTCAAGGAAGCCCAACAGTCGCGCTCTTAAATAATACCGACTTCCAA 628
 |||||
 Oy 1382 CAGCAGCATCTTACGAGCCAGCAAAATTAAGATGCTCAGTATTAATAATATTTT 1441
 |||||
 Db 629 CAGCAGCATCTTACGAGCCAGCAAAATTAAGATGCTCAGTATTAATAATATTTT 688
 |||||
 Oy 1442 CC 1443
 |||||
 Db 689 CC 690
 |||||
 RESULT 2
 AV025152
 LOCUS
 DEFINITION
 AV025152 Mus musculus adult C57BL/6J lung Mus musculus cDNA clone
 1200010P16, mRNA sequence.
 ACCESSION
 AV025152
 VERSION
 AV025152.2 GI:154311435
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 721)
 REFERENCE
 AUTHORS
 Aizawa, K., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
 Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
 Kono, H., Konda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
 Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Aizawa, T., et al. 2001)
 Unpublished (2001)
 TITILE
 JOURNAL
 COMMENT
 On May 11, 1999 this sequence version replaced gi:4802144.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@rken.go.jp/
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A.
 and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamana, K.,
 Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K.,
 Aizawa, K., Ishii, Y., and Hayashizaki, Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct.*

Func. Genomics 2 pre, 172-L86 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.

FEATURES

Source

Location/Qualifiers
1. 721
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="120010P16"
/issue_type="lung"
/dev_stage="adult"
/clone_lib="Mus musculus adult C57BL/6J lung"

ORIGIN

Query Match 35.7%; Score 557; DB 9; Length 721;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 607; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 876 TCCAAATTTTACAGATCAGAAAGCATCCGGAAGAGCTTACGTCAGCTGGCATAC 935
DB 61 TCCAAATTTTACAGATCAGAAAGCATCCGGAAGAGCTTACGTCAGCTGGCATAC 120
QY 936 TCTGAGAGCAGCAGATCCGAGTGTCCCAACAAGACTACGACTCTCCGACGCTTTC 995
DB 121 TCTGAGAGCAGCAGATCCGAGTGTCCCAACAAGACTACGACTCTCCGACGCTTTC 180
QY 996 CTGCTCATGTTTCTTCTTCATCATGTGAGATGCCATCATCATACCATCTCTCATC 1055
DB 181 CTGCTCATGTTTCTTCTTCATCATGTGAGATGCCATCATCATACCATCTCTCATC 240
QY 1056 TTGATCCAAATCTCCGAGAGCATCTGTGATCTGGCCATCCCTTTCTTGTGGTGTG 1115
DB 241 TTGATCCAAATCTCCGAGAGCATCTGTGATCTGGCCATCCCTTTCTTGTGGTGTG 300
QY 1116 GCCTTCAGTTTGCACATCTGCCTTAACCCCATACCTGATCATGTGCTGTGAGG 1175
DB 301 GCCTTCAGTTTGCACATCTGCCTTAACCCCATACCTGATCATGTGCTGTGAGG 360
QY 1176 AACGAAATGAGAGAAATTTTGTGCTTCTTTTTCAGAGAGGAGCCATTTTACA 1235
DB 361 AACGAAATGAGAGAAATTTTGTGCTTCTTTTTCAGAGAGGAGCCATTTTACA 420
QY 1236 GATACGTTCTGACGGGAATGACTTGTCTGTATTTCCAGCTAACTAGCTCTGTGCC 1295
DB 421 GATACGTTCTGACGGGAATGACTTGTCTGTATTTCCAGCTAACTAGCTCTGTGCC 480
QY 1296 AGGTGAACACGGGTGTGATGTAAGGGAGTTAACTCAAGAAAGCCACAGTGGCC 1355
DB 481 AGGTGAACACGGGTGTGATGTAAGGGAGTTAACTCAAGAAAGCCACAGTGGCC 540
QY 1356 CTGCTTTAAATATACCGACTTCCACAGAGGAGCTTACGAGCCAGCAATTAAGAA 1415
DB 541 CTGCTTTAAATATACCGACTTCCACAGAGGAGCTTACGAGCCAGCAATTAAGAA 600
QY 1416 TGAATCGCTGATTAATAATTTTCTTAAAGAACTTTCTATGGGTTCTTTTGTGA 1475
DB 601 TGAATCGCTGATTAATAATTTTCTTAAAGAACTTTCTATGGGTTCTTTTGTGA 660
QY 1476 ACTTTT 1483
DB 661 ACTTTT 668

```

RESULT 3
BO939003 927 bp mRNA linear EST 21-AUG-2002
LOCUS BO939003
DEFINITION AGENCOURT_8946852 NCI_CGAP_Co24 Mus musculus cDNA clone

IMAGE:6395027 5', mRNA sequence.

ACCESSION BO939003
VERSION BO939003.1 GI:22354481

KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerogomathi; Muridae; Murinae; Mus.

REFERENCE

NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LHAM13889 Row: a Column: 12
High quality sequence stop: 624.

FEATURES

Source

Location/Qualifiers
1. 927
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6395027"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; Not1; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 33.7%; Score 525; DB 13; Length 927;

Best Local Similarity 99.5%; Pred. No. 0;
Matches 725; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 231 CCTTCGACACCTCGACCAAGTCAATGCAACCACTTCTTCTGGAATCAAG 290
DB 52 CCTTCGACACCTCGACCAAGTCAATGCAACCACTTCTTCTGGAATCAAG 111
QY 291 GGCACACACCGGTGGTGGTGAAGGTGTGAGAGCAACCGTTCTGGAGTCACTTTGTC 350
DB 112 GGCACACACCGGTGGTGGTGAAGGTGTGAGAGCAACCGTTCTGGAGTCACTTTGTC 171
QY 351 GTCTCACTGTGGGCAACGTGTGCTCTAGTCTGTGTGCGCGCGTGGCGCGTGGG 410
DB 172 GTCTCACTGTGGGCAACGTGTGCTCTAGTCTGTGTGCGCGCGTGGCGCGTGGG 231
QY 411 GGTCAACGACGCTGTGTCTCAACCTTTCTGGCGGATTTGCTTCCACAGGCGCATC 470
DB 232 GGTCAACGACGCTGTGTCTCAACCTTTCTGGCGGATTTGCTTCCACAGGCGCATC 291
QY 471 CTCTAGTGTCTGTGCGCTGAGTGAAGGCTGTGTTTGGGAGCCGCTGTGTGCAC 530
DB 292 CTCTAGTGTCTGTGCGCTGAGTGAAGGCTGTGTTTGGGAGCCGCTGTGTGCAC 351
QY 531 CTGCTTTCTAGTATGATGACAAATGAGCGGAGCGTCAAGTCTTCACACTGGCCGCTC 590
DB 352 CTGCTTTCTAGTATGATGACAAATGAGCGGAGCGTCAAGTCTTCACACTGGCCGCTC 411
QY 591 AGCTGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 650
DB 412 AGCTGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 471
QY 651 CGGACTGAGGAGGAGTGTGCTTTCTATATGAGGTTACTCGGCGCTCGGCGCTGCCC 710
DB 472 CGGACTGAGGAGGAGTGTGCTTTCTATATGAGGTTACTCGGCGCTCGGCGCTGCCC 531

```

Qy	711	CTCTACATCTTGTTCCGGGTGTTCCCGAGAGCCCTCCCGCGGGGACACAGAAATTCG	770
Db	532	CTTCGCATCTTGTTCCGGGTGTTCCCGAGAGCCCTCCCGCGGGGACACAGAAATTCG	591
Qy	771	ATTTCACATTTGATTGGCCCAACCGCATAGAGAAATCTATGGATGTGTTTTTGAG	830
Db	592	ATTTCACATTTGATTGGCCCAACCGCATAGAGAAATCTATGGATGTGTTTTTGAG	651
Qy	831	ACTTTGAACCTTCCTGTGTCCCGGAACTGGTCAATTGTGATCACTTACTCCAAAATTTTACAG	890
Db	652	ACTTTGAACCTTCCTGTGTCCCGGAACTGGTCAATTGTGATCACTTACTCCAAAATTTTACAG	711
Qy	891	ATTCAGAAAGCATCGCGGAGAGAGGCTTACGTGAGCTTGAGCATACTGAGAGCCACCAG	950
Db	712	ATTCAGAAAGCATCGCGGAGAGAGGCTTACGTGAGCTTGAGCATACTGAGAGCCACCAG	771
Qy	951	ATCCGAGTG 959	
Db	772	ATCCGAGTG 780	

RESULT 4	
BUS22456	
LOCUS	
DEFINITION	BUS22456 870 bp mRNA linear EST 13-SEP-2002
	AGNCCOURT_10153856 NCI CGAL_CQ24 Mus musculus cDNA clone
	IMAGE:6528871 5', mRNA sequence.

ACCESSION	BUS22456	GI:2282982
VERSION	BUS22456.1	
KEYWORDS	EST.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (pages 1 to 870)	NH-MGC	http://mgc.ncl.nih.gov/ .	
Unpublished (1999)	National Institutes of Health,	Mammalian Gene Collection (MGC)	
Contact:	Robert Strausberg,	Ph.D.	

Email: cgapbbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LMNL128 row: b column: 07
High quality sequence stop: 615.
Location/Qualifiers
1..870

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6528871"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1fb="NCI CGAP Co24"
/notes="Organ: colon, Vector: pCMV-SPORT6, Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by life
technologies. Note: this is a NCI_CGAP library."

```

Query Match	30.8%	Score 481;	DB 13;	Length 870;
Best Local Similarity	99.4%	Pred. No. 0;		
Matches 681; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0

Oy 231 CCTGTGCACACCCCTGGACCAAGTCAATGGCAACCACTTCCCTTTCTTCCGATGTCAAG 290
 Db 48 CCTGTGCACACCCCTGGACCAAGTCAATGGCAACCACTTCCCTTTCTTCCGATGTCAAG 107
 Oy 291 GCGCAGCAACCGGTGTGTGAGCGTCTGTGAGCAACCGTTCTTGGAATCATCTTTGTC 350

Db	108	GGCGACACACCGGTTGGTGTGTGTTAGGCTGTGTGAGACACACCGTTCTGGGGGCTCATCTTTGTCT	167
Qy	351	GTCTCACTGCTGGGCAACGTGTGTGCTCTAGTGTGTGTGGCGCGCCGTGGCGGCGCTGGGG	410
Db	168	GTCTCACTGCTGGGCAACGTGTGTGCTCTAGTGTGTGTGGCGCGCCGTGGCGGCGCTGGGG	227
Qy	411	GGGTGACGCGAGCCGTGGTGTCTCAACCTCTTACGGCGGATTTGGCTCTTCAACGAGCGCATC	470
Db	228	GGGAGCGACCGCTGGTGTCTTCAACCTCTTCTGGCGGATTTGGCTCTTCAACGAGCGCATC	287
Qy	471	CCTCTAGTGTCTGT	530
Db	288	CCTCTAGTGTCTGT	347
Qy	531	CTGTCTTCTTACGT	590
Db	348	CTGTCTTCTTACGT	407
Qy	591	AGCCTGAGGCGCATGT	650
Db	408	AGCCTGAGGCGCATGT	467
Qy	651	CGGACTACGGGGGCACTGT	710
Db	468	CGGACTACGGGGGCACTGT	527
Qy	711	CTTCAACATCTTGT	770
Db	528	CTTGT	587
Qy	771	ATTGTGACATTTGT	830
Db	588	ATTGTGACATTTGT	647
Qy	831	ACTTTGAACCTTCTGT	890
Db	648	ACTTTGAACCTTCTGT	707
Qy	891	ATCAAGAAAGCATCGCGGAAGAGCG 915	
Db	708	ATCAAGAAAGCATCGCGGAAGAGCG 732	

RESULT 5					
BUS522608					
LOCUS					
DEFINITION	BUS522608	875 bp	mRNA	linear	EST 13-SEP-2002
	AGNCNCOURT_10154009	NCI CGAP Co24	Mus musculus	cDNA clone	
IMAGE:6529120	5',	mRNA sequence.			

ACCESSION	BUS22608	
VERSION	BUS22608.1	GI:22830134
KEYWORDS	EST.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 875)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strusberg, Ph.D.

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Email: cs9abs-r@mail.nih.gov
Tissue Procurement: The Cepco Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L14M14128 row: 1 column: 16
High quality sequence stop: 653.

FEATURES	Location/Qualifiers
SOURCE	1. .875

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6529120"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Co24"
/note="Organ: colon; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI CGAP library."

```

```

KEYWORDS EST
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. (bases 1 to 928)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LAM4034 row: b column: 04 High quality sequence, stop: 605.

```

```

Query Match 29.8%; Score 465; DB 13; Length 875;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 665; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

FEATURES
source
1..928
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6488067"
/issue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam2"
/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; SalI; Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

```

```

231 CCTCGGACACCCGAGCAAGTCAATGACACCACTTCCCTTCTTCTGAGATGTCAG 290
84 CCTCGGACACCCGAGCAAGTCAATGACACCACTTCCCTTCTTCTGAGATGTCAG 143
291 GGGGACACCCGAGTGTGTGAGCGTGTGAGACCAACCGTTCTGGGACTCATCTTGTG 350
144 GGGGACACCCGAGTGTGTGAGCGTGTGAGACCAACCGTTCTGGGACTCATCTTGTG 203
351 GTCTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 410
204 GTCTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 263
411 GCGTCAAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470
264 GCGTCAAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 323
471 CCTTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 530
324 CCTTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 383

```

```

Query Match 28.5%; Score 445; DB 13; Length 928;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 595; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

531 CTGCTCTTTCAAGTGAATGACATGAGGCGGACGCTCAACATCTTCAACATGAGCGG 590
384 CTGCTCTTTCAAGTGAATGACATGAGGCGGACGCTCAACATCTTCAACATGAGCGG 443
591 AGCCTGAGGCGCATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 650
444 AGCCTGAGGCGCATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 503
651 CGGACTCAGGCGGACGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 710
504 CGGACTCAGGCGGACGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 563
711 CTCTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 770
564 CTCTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 623
771 ATTGACATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 830
624 ATTGACATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 683
831 ACTTTGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 890
684 ACTTTGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
891 ATCAGGAAA 899
744 ATCAGGAAA 752

```

```

231 CCTCGGACACCCGAGCAAGTCAATGACACCACTTCCCTTCTTCTGAGATGTCAG 290
69 CCTCGGACACCCGAGCAAGTCAATGACACCACTTCCCTTCTTCTGAGATGTCAG 128
291 GGGGACACCCGAGTGTGTGAGCGTGTGAGACCAACCGTTCTGGGACTCATCTTGTG 350
129 GGGGACACCCGAGTGTGTGAGCGTGTGAGACCAACCGTTCTGGGACTCATCTTGTG 188
351 GTCTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 410
189 GTCTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 248
411 GCGTCAAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470
249 GCGTCAAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 308
471 CCTTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 530
309 CCTTCACTGCTGGGCAACGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 368
531 CTGCTCTTTCAAGTGAATGACATGAGGCGGACGCTCAACATCTTCAACATGAGCGG 590
369 CTGCTCTTTCAAGTGAATGACATGAGGCGGACGCTCAACATCTTCAACATGAGCGG 428
591 AGCCTGAGGCGCATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 650
429 AGCCTGAGGCGCATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 488
651 CGGACTCAGGCGGACGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 710
489 CGGACTCAGGCGGACGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCT 548

```

```

RESULT 6
LOCUS B0936343 928 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_10013769 NCI CGAP Mam2 Mus musculus cDNA clone
IMAGE:6488067 5', mRNA sequence.
ACCESSION B0936343
VERSION B0936343.1 GI:22351726

```

QY 711 CTCATCATCTTGTTCCTCCGCTGCTCCCGACGCGCTTCCTCCGCGGAGACAGAAATTCG 770
 DB 549 CTCGATCTTGTTCCTCCGCTGCTCCCGACGCGCTTCCTCCGCGGAGACAGAAATTCG 608
 QY 771 ATTTGCATTGATTGGCCCAACCGCATMGAGAAATCTCATGGAGTGTGTTTTCG 828
 DB 609 ATTGACATTGATTGGCCCAACCGCATMGAGAAATCTCATGGAGTGTGTTTTCG 666

RESULT 7
 BX527210 537 bp mRNA linear EST 27-JUN-2003
 LOCUS BX527210 Soares mammary gland NbMmg Mus musculus cDNA clone
 DEFINITION IMA6P98K072670 ; IMAGE:1067646, mRNA sequence.
 ACCESSION BX527210
 VERSION BX527210.1 GI:32304990
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 537)
 Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.
 and Korn,B.
 Mouse Unigeneset - RZPD2
 Unpublished (2003)
 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMA6P98K072670.
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
 http://www.rzpd.de/Cloncard/cg1-
 bin/showlib.pl.cgi?response?libNo=981 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de

FEATURES
 source
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 1T, Primer sequence: TAAATGCACTCACTATAGGG.

Location/Qualifiers
 1..537
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMA6P98K072670 ; IMAGE:1067646"
 /sex="male"
 /library_type="mammary gland"
 /issue_type="4 weeks"
 /dev_stage="DHI0B"
 /lab_host="DHI0B"
 /clone_id="Soares mammary gland NbMmg"
 /note="Organ: mammary gland; Vector: pRTT3D-Pac
 (pharmacia) with a modified polylinker; Site 1: Not I,
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5,
 TGTTACCAATCTGATGAGTGAGGCGCGCGCAAGGTGTTTGTGTTTGTGTTTGTGTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pRTT3 vector.
 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
 constructed and normalized by Bento Soares and M.Fatima
 Bernaldo."

ORIGIN

Query Match 27.8%; Score 434; DB 13; Length 537;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 533; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 561 AGGTCACGATCTTCACATCTGCGCGGCTGAGCGCATGTGTGATCTGCGC 620
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 2 AGGTCACGATCTTCACATCTGCGCGGCTGAGCGCATGTGTGATCTGCGC 61
 QY 621 CTCGCGCGCGGCTTGAAGCGGCGCGGCGGCGGACTCAGCGGCACTGTGCTTTCATA 680
 DB 62 CTCGCGCGCGGCTTGAAGCGGCGCGGCGGCGGCACTCAGCGGCACTGTGCTTTCATA 121
 QY 681 TGGGGTTACTCGGCGCTCGCGGCTGCGGCTCTTCACTCTTGTTCGCGTCCGCGAG 740
 DB 122 TGGGGTTACTCGGCGCTCGCGGCTGCGGCTCTTCACTCTTGTTCGCGTCCGCGAG 181
 QY 741 CGGCTTCCGCGGCGGAGACAGAGAAATTCGATTTGACATTTGATTTGGCCCAACCGGATA 800
 DB 182 CGGCTTCCGCGGCGGAGACAGAGAAATTCGATTTGACATTTGATTTGGCCCAACCGGATA 241
 QY 801 GGAGAAATCTCATGAGATGTGTTTGTGAGACTTTGAACCTTCCGTGCGGAGACTGCTC 860
 DB 242 GGAGAAATCTCATGAGATGTGTTTGTGAGACTTTGAACCTTCCGTGCGGAGACTGCTC 301
 QY 861 ATTGTGATCAGTTACTTCCAAATTTTACAGATCAGAGACATCGCGGAGAGGCTTACG 920
 DB 302 ATTGTGATCAGTTACTTCCAAATTTTACAGATCAGAGACATCGCGGAGAGGCTTACG 361
 QY 921 CTGAGCTTGGCATATCTTGAGAGCCACAGATCCGAGTGTCCCAACAGACTACCGGCTC 980
 DB 362 CTGAGCTTGGCATATCTTGAGAGCCACAGATCCGAGTGTCCCAACAGACTACCGGCTC 421
 QY 981 TTCCGACGCTTCTCTGCTCATGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1040
 DB 422 TTCCGACGCTTCTCTGCTCATGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 481
 QY 1041 ACCATCTCTCTCATGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1096
 DB 482 ACCATCTCTCTCATGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 537

RESULT 8
 LOCUS BB609925
 DEFINITION BB609925 RIKEN full-length enriched, adult male lung Mus musculus
 cDNA clone 1200010P16 5', mRNA sequence.
 ACCESSION BB609925
 VERSION BB609925.1 GI:15390516
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 671)
 Arikawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,D.,
 Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,H., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tgawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
 Riken Mouse ESTs (Arikawa,T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayashizaki,Y., Sugahara,Y., Shibata,K.,
 Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Matshiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,

	TITLE	JOURNAL	COMMENT
Oy	196 GATCCGAGTGGCCCCAACAAGAAGCTACCGACTCTTCGCCAGCGCTCTTCGTGCATGATTGC	255	
Oy	1010 CTTCTTATCATATGATGGAGTGCCCATCATCATCATACATCTCTCTCATCTTTGATCCAAACTT	1069	
Db	256 CTTCTTATCATATGATGGAGTGCCCATCATCATCATCTCTCTCATCTTTGAATCAAAACTT	315	
Oy	1070 CCGCAGAGAACCTGGTATCTGTCGGCCATCCCTTTTCTTCGAGGTGGTGCCCTCACGTTGC	1129	
Db	316 CCGCAGAGAACCTGGTATCTGTCGGCCATCCCTTTTCTTCGAGGTGGTGCCCTCACGTTGC	375	
Oy	1130 CAACCTGACCCTTAACCCCATCATCTGTACACATGTGCTGTTCAGGAACGAATGAGGAA	1189	
Db	376 CAACCTGACCCTTAACCCCATCATCTGTACACATGTGCTGTTCAGGAACGAATGAGGAA	435	
Oy	1190 GATTTTTTGCTGCTCTTTCTTTTCCAGAGAAAGGAGCCATTTTACAGA	1237	
Db	436 GATTTTTTGCTGCTCTTTCTTTTCCAGAGAAAGGAGCCATTTTACAGA	483	
RESULT 10			
LOCUS	B609814	637 bp	mRNA linear EST 26-OCT-2001
DEFINITION	B609814 RIKEN full-length enriched, 18 days embryo Mus musculus		
ACCESSION	CNA clone 1190023B03 5', mRNA sequence.		
VERSION	B609814		
KEYWORDS	B609814.1 GI:16451529		
SOURCE	EST.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus.		
	1 (bases 1 to 637)		
	Atakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hizumoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,U., Komori,H., Kouda,M., Koya,S., Matsumura,T., Miyazaki,A., Nomura,K., Ono,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shingawa,A., Shiraki,T., Sugabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,W. and Hayashizaki,Y.		
	RIKEN Mouse ESTs (Atakawa,T., et al. 2001)		
	Unpublished (2001)		
	Contact: Yoshihide Hayashizaki		
	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute		
	The Institute of Physical and Chemical Research (RIKEN)		
	1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan		
	Tel: 81-45-503-9222		
	Fax: 81-45-503-9216		
	Email: genome-res@gsr.riken.go.jp,		
	URL:http://genome-gsc.riken.go.jp/,		
	Carninci,P., Shibata,Y., Hayasaka,N., Sugahara,Y., Shibata,K., Itoh,M., Komori,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)		
	wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuhiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)		
	Komori,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.		
	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)		
	Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamataka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.		
	Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome 12, 673-677 (2001)		
	Please visit our web site (http://genome-gsc.riken.go.jp/) for		

FEATURES	Location/Qualifiers
source	1. 637
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	/clone="1190023B03"
	/dev_stage="18 days embryo"
	/lab_host="SOLR"
	/clone_1db="RIKEN full-length enriched, 18 days embryo"
	/note="Site_1: XhoI; Site_2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5],
	GAAGAGAGAGCGCGCGCAACTCGAGTCTTTTTTTTTTTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5]
	GAAGAGAGAGAGATCCAGAGCTCAATTATTAATTAATAACCCCCCCC 3'.
	cDNA was cleaved with XhoI and SstI."

ORIGIN	Query Match	24.5%; Score 382; DB 10; Length 637;
Best Local Similarity 99.6%; Pred. No. 0;		
Matches 482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	231	CCCTGCGACACCCCTGGACCAAGTCATCGACCCACTTCCCTTTCTTCGAGATGCAAG 290
Db	68	CCCTGCGACACCCCTGGACCAAGTCATCGACCCACTTCCCTTTCTTCGAGATGCAAG 127
QY	291	GCGACACACCGGTTGGTGTGAGCGTCTGAGAGCACCGTTCTGGACTCATTTTGTG 350
Db	128	GCGACACACCGGTTGGTGTGAGCGTCTGAGAGCACCGTTCTGGAGGCTCATTTTGTG 187
QY	351	GTTCTACTGCTGGGCAACGTGTGTCTTATGTGCTGTGGCGCGCCGTGGCGCCGTGGG 410
Db	188	GTTCTACTGCTGGGCAACGTGTGTCTTATGTGCTGTGGCGCGCCGTGGCGCCGTGGG 247
QY	411	GCGTCACCGACCGCTGGGTCCTCAACCTTTCGCGCGGATTGTCCTTCACAGCGGCATC 470
Db	248	GCGTCACCGACCGCTGGGTCCTCAACCTTTCGCGCGGATTGTCCTTCACAGCGGCATC 307
QY	471	CCTTATGTGCTCGTCGTCGCTGAGACTGAGCGCTGCTGTGGGCGCCGTGTCGTCAC 530
Db	308	CCTTATGTGCTCGTCGTCGCTGAGACTGAGCGCTGCTGTGGGCGCCGTGTCGTCAC 367
QY	531	CTGCTCTTTCACGTGATGACATATGAGCGGACGCTCACGATCCTCACATGCGCGGTC 590
Db	368	CTGCTCTTTCACGTGATGACATATGAGCGGACGCTCACGATCCTCACATGCGCGGTC 427
QY	591	AGCTTGAAGGCGCATGCTGTCATCTGTCGCGCTCCGCGCGCGCTTGAAGCGCGCGGCGG 650
Db	428	AGCTTGAAGGCGCATGCTGTCATCTGTCGCGCTCCGCGCGCGCTTGAAGCGCGCGGCGG 487
QY	651	CGGACTGAGGCGGCACTGCTGCTTTCATATGAGGGTTACTCGGCGCTCGCGCGCTGCC 710
Db	488	CGGACTGAGGCGGCACTGCTGCTTTCATATGAGGGTTACTCGGCGCTCGCGCGCTGCC 547
QY	711	CTCT 714
Db	548	CTCT 551

RESULT 11	A1552415/c	437 bp	mRNA	linear	EST 23-MAR-1999
LOCUS	A1552415	437 bp	mRNA	linear	EST 23-MAR-1999
DEFINITION	VP04e04 x1 Soares_mammary_gland NbMMG Mus musculus cDNA clone IMAGE:1067646 3', mRNA sequence.				

ACCESSION	A1552415
VERSION	A1552415.1
KEYWORDS	GI:4484778
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
JOURNAL COMMENT	Mus musculus Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCI-CCAF http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgebps-xremail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGJ:59006 This clone was previously sequenced on the 5' end only, this new data is from the 3' end High quality sequence stop: 399.
REFERENCE AUTHORS TITLE	1 (bases 1 to 437)
FEATURES SOURCE	Location/Qualifiers 1..437 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6J" /db_xref="taxon:10090" /clone="IMAGE:1067646" /sex="male" /tissue_type="mammary gland" /dev_stage="4 weeks" /lab_host="DH10B" /clone_lib="Soares mammary gland NbmWG" /note="Organ: mammary gland; Vector: pTY73D-Pac (pharmacia) with a modified polylinker; Site 1: Not I, Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5' TGTTACCATCTCAGATGAGGAGCGGCAGATGTATTTTTTTTTTTTTTTTTTTT T 3'] : double-stranded cDNA was ligated to Eco RI adaptors (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTY73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
ORIGIN	
Query Match	17.8%; Score 278; DB 9, Length 437;
Best Local Similarity	99.5%; Pred. No. 0;
Matches 378; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1104 TTCTGGGGTGGGGGCTTCAGCTTTGGGCCAACTCTGCGCTTAACCATTCTGTACAACATG 1163
DB	437 TTTCTGGGGTGGGGCTTCAGCTTTGGGCCAACTCTGCGCTTAACCATTCTGTACAACATG 378
OY	1164 TCCTGTGTTCAGGAAGAATGAGAGAAATTTTTTCTGCTCTTTTTCAGAGAAGGA 1223
DB	377 TCCTGTGTTCAGGAAGAATGAGAGAAATTTTTTCTGCTCTTTTTCAGAGAAGGA 318
OY	1224 GCCATTTTTCAGATACGTCTGTCAAGCGGAATGACTTGTCTGTTATTTCCAGCTACTA 1283
DB	317 GCCATTTTTCAGACACGCTCTGTCTCAGCGGAATGACTTGTCTGTTATTTCCAGCTACTA 258
OY	1284 GCCTCTGTGTTCAGGTGAACACGCGTGTGCATGTAAAGGAGTTAACTTCAAAGGAAGCC 1343
DB	257 GCCTCTGTGTTCAGGTGAACACGCGTGTGCATGTAAAGGAGTTAACTTCAAAGGAAGCC 198
OY	1344 CACCAAGTCGCGCTCTTTAAAATAACCGACTTCCAACAGAGCATCTACGAGCCAG 1403
DB	197 CACCAAGTCGCGCTCTTTAAAATAACCGACTTCCAACAGAGCATCTACGAGCCAG 138
OY	1404 CAATTAAGGATGATCGCTCAGTTAAATAATTTTCTTTAAAGAACTTTCTATGGG 1463
DB	137 CAATTAAGGATGATCGCTCAGTTAAATAATTTTCTTTAAAGAACTTTCTATGGG 78

QY	1464	TTCCCTTTGACACTTTT	1463
Db	77	TTCCCTTTGACACTTTT	58
RESULT 12			
LOCUS	CG617050	492 bp	DNA
DEFINITION	OSJ310062 Mus musculus 129Sv/Ev Mus musculus genomic clone		
ACCESSION	CG617050		
VERSION	CG617050.1	GI:37440899	
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 492)		
	Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Piggett, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A., Frixidde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C., Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.		
TITLE	Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)		
COMMENT	Contact: Zambrowicz BP		
	OmlBank		
	Lexicon Genetics Incorporated		
	4000 Research Forest Drive, The Woodlands, TX 77381, USA		
	Email: materials@lexgen.com		
	Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)		
	Class: Gene Trap.		
FEATURES			
source	Location/Qualifiers		
	1..492		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/strain="129Sv/Ev"		
	/db_xref="taxon:10090"		
	/clone="OSTJ10062"		
	/cell_type="embryonic stem cell"		
	/clone_id="Mus musculus 129Sv/Ev"		
ORIGIN			
Query Match	17.2%	Score 269;	DB 29; Length 492;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 269;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	969	GACTACGACCTTCGCGACGCTTCCTCGTGATCGATTCCTTCATCATGTGAGT	1028
Db	173	GACTACGACCTTCGCGACGCTTCCTCGTGATCGATTCCTTCATCATGTGAGT	232
QY	1029	CCCATCATCATCACCATCTCTCTCATCTTGATCCAAACTTCGCGACGACCTGTATC	1088
Db	233	CCCATCATCATCACCATCTCTCTCATCTTGATCCAAACTTCGCGACGACCTGTATC	292
QY	1089	TGGCCATCCCTTTTCTTCTGGGTGTGTGCGCTTACCGTTTGCAACTGTGCGCTTAAACCC	1148
Db	293	TGGCCATCCCTTTTCTTCTGGGTGTGTGCGCTTACCGTTTGCAACTGTGCGCTTAAACCC	352
QY	1149	ATATGTATCAACATGTGCTGTTCAGGAAGCAATGAGAGAAATTTTGTGCTGCTCTT	1208
Db	353	ATATGTATCAACATGTGCTGTTCAGGAAGCAATGAGAGAAATTTTGTGCTGCTCTT	412
QY	1209	TTTCAGAGAGAGGACCATTTTTCAGA 1237	
Db	413	TTTCAGAGAGAGGACCATTTTTCAGA 441	
RESULT 13			

BF580577 647 bp mRNA linear EST 12-DEC-2000
 LOCUS 602097386F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4217252 5',
 DEFINITION mRNA sequence.
 ACCESSION BF580577
 VERSION BF580577.1 GI:11654289
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 647)
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strauberg, Ph.D.
 COMMENT Email: cgapds-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.lnl.gov
 Plate: LM9795 row: 1 column: 21
 High quality sequence stop: 626.
 Location/Qualifiers
 1..647
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4217252"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_id="NCI_CGAP_Co24"
 /note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
 ORIGIN
 Query Match 13.7%; Score 214; DB 10; Length 647;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 913 GGCCTAGCGGAGCTTGGCATCTGAGAGCCAGATCCGAGTCCCAAGAAGCT 972
 Db 1 GGCCTAGCGGAGCTTGGCATCTGAGAGCCAGATCCGAGTCCCAAGAAGCT 60
 QY 973 ACCGACTCTCCGACGCTCTTCTGCTCATGTTCCTTCTTCATCATGTGAGTCCCA 1032
 Db 61 ACCGACTCTCCGACGCTCTTCTGCTCATGTTCCTTCTTCATCATGTGAGTCCCA 120
 QY 1033 TCATCATCACCATCTCTCTCATCTTGAATCCAAAATCTCCGACGAGCTGTCTATCTGGC 1092
 Db 121 TCATCATCACCATCTCTCTCATCTTGAATCCAAAATCTCCGACGAGCTGTCTATCTGGC 180
 QY 1093 CATCCCTTTCTTCTGGGTGGTGGCTTCAAGTT 1126
 Db 181 CATCCCTTTCTTCTGGGTGGTGGCTTCAAGTT 214
 RESULT 14
 LOCUS AA030752 365 bp mRNA linear EST 21-AUG-1996
 DEFINITION m31h04.t1 Soares mouse embryo NDMB13.5 14.5 Mus musculus cDNA
 clone IMAGE:465175 5', mRNA sequence.
 ACCESSION AA030752
 VERSION AA030752.1 GI:1500749
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 365)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellendberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LMNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:278991
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 356.
 Location/Qualifiers
 1..365
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:465175"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I oligo(7nt) primer [5', TGTTACCAATCTAAGTGGAGGCGCGGGAATTTTCTTTTCTTTTCTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773D vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Patima Bonaldo. "
 ORIGIN
 Query Match 12.4%; Score 194; DB 9; Length 365;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 364; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 904 CGGGAGAGAGGCTTACGCTGAGCTGAGCATCTGAGAGCCAGATCCGAGTCCC 963
 Db 1 CGGGAGAGAGGCTTACGCTGAGCTGAGCATCTGAGAGCCAGATCCGAGTCCC 60
 QY 964 AACAAAGTACCGACTCTTCGCGACGCTCTTCGCTCATGTTCCTTCTTCATCATGT 1023
 Db 1 AACAAAGTACCGACTCTTCGCGACGCTCTTCGCTCATGTTCCTTCTTCATCATGT 120
 QY 964 AACAAAGTACCGACTCTTCGCGACGCTCTTCGCTCATGTTCCTTCTTCATCATGT 1023
 Db 61 AACAAAGTACCGACTCTTCGCGACGCTCTTCGCTCATGTTCCTTCTTCATCATGT 120
 QY 1024 GGAATCCCATCATCATCATCATCTCTCATCTTGAATCCAAAATCTCCGACGAGACCTGG 1083
 Db 121 GGAATCCCATCATCATCATCATCTCTCATCTTGAATCCAAAATCTCCGACGAGCTGG 179
 QY 1084 TCATCTGGCCATCCCTTTCTTCTGGGTGGTGGCTTCAAGTTTGCACATCTGCCCCAA 1143
 Db 180 TCATCTGGCCATCCCTTTCTTCTGGGTGGTGGCTTCAAGTTTGCACATCTGCCCCAA 239
 QY 1144 ACCCCATACGTGTACAACATGTGCGTGTTCAGAAAGAAAGAGAGATTTTGTGCT 1203
 Db 240 ACCCCATACGTGTACAACATGTGCGTGTTCAGAAAGAAAGATTTTGTGCT 299
 QY 1204 TCTTTTTCAGAGAGAGGAGCCATTTTTCAGATACGTCTGTACGCGGAATATCTTGT 1263

Db 300 TCCTGTTTCCAGAGAGGAGCATTTTACAGATAGCTGTCGAGGGAATGACTTGT 359
QY 1264 CTGTTA 1269
Db 360 CTGTTA 365

RESULT 15
BY345055 344 bp mRNA linear EST 12-DEC-2002
LOCUS BY345055 RIKEN full-length enriched, whole joints Mus musculus cDNA
DEFINITION clone U730011C21 5', mRNA sequence.
ACCESSION BY345055
VERSION BY345055.1 GI:26574543
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 344)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
Batilov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
Chochua, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gasteier, T.,
Geroldi, M., Gissi, C., Godzik, A., Gough, J., Grimmer, S.,
Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kani, A.,
Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kongaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Maltsev, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Nimata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. V., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandalin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
Verardo, L., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayashizaki, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Atawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE JOURNAL MEDLINE PUBMED
COMMENT 12466851

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Aizawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapped-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
Fleming street 16672 Vari, Greece) whose assistance we gratefully
acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1. 344
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="U730011C21"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN
Query Match 11.5%; Score 180; DB 13; Length 344;
Best Local Similarity 99.3%; Pred. NO. 0;
Matches 280; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
231 CCCGCGACACCCCTGGACCAAGTCAATGCAACCCACTTCCCTTCTCGAGATGCAAG 290
Db CCCGCGACACCCCTGGACCAAGTCAATGCAACCCACTTCCCTTCTCGAGATGCAAG 122
QY GGGACACACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 350
Db GGGACACACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 182
QY GTCTCACTGCTGGGCAACGTGTGCTCTAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 410
Db GTCTCACTGCTGGGCAACGTGTGCTCTAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 242
QY GCGCGACGACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 470
Db GCGCGACGACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 302
QY CCTCTAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 512
Db CCTCTAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 344

Search completed: October 1, 2004, 03:51:49
Job time : 3677.59 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 19:03:02 ; Search time 6615.8 Seconds

(without alignments)
11419.163 Million cell updates/sec

Title: US-10-077-698-2

Perfect score: 1743

Sequence: 1 tccgaccagctctcagccg.....aaaaaaaaaagggcgg 1743

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

```
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1743	100.0	1743	6	BD240720 14273 rec
2	1743	100.0	1743	6	BD269630 14273 rec
3	1743	100.0	1743	6	AR228216 Sequence
4	1743	100.0	1743	6	AR372101 Sequence
5	1328	76.2	1458	6	BD015337 Novel pol
6	1081	62.0	1321	6	AX168141 Sequence
7	986	56.6	163964	9	AX356214 Human DNA
8	926	53.1	1086	6	AX498184 Sequence
9	693	39.8	744	9	AY255573 Homo sapi
10	559	32.1	1066	6	AX647273 Sequence
11	445	25.5	1131	9	AY288417 Homo sapi
12	425	24.4	933	6	AX657504 Sequence
13	391	22.4	1737	6	AX549396 Sequence
14	326	18.7	1104	6	AX230123 Sequence
15	300	17.2	300	6	BD212934 Novel hum
16	298	17.1	760	6	BD129472 Human gen
17	198	11.4	300	6	BD128746 Human gen
18	143	8.2	300	6	BD128747 Human gen
19	131	7.5	638	6	AX277776 Sequence
20	131	7.5	4291	9	BSM607867 Homo sapi
21	111	6.4	426	6	AX230146 Sequence
22	68	3.9	39915	9	AC010512 Homo sapi
23	67	3.8	100300	2	AC114970 Homo sapi
24	67	3.8	110000	2	AC139803_1
25	67	3.8	110000	2	AC140020_0
26	67	3.8	129218	2	AC093280
27	67	3.8	129218	2	AC093851
28	67	3.8	129218	2	AC138851
29	67	3.8	136785	2	AC139460
30	67	3.8	138343	2	AC139273
31	67	3.8	139121	2	AL135916
32	67	3.8	142834	2	AC139499
33	67	3.8	145493	2	AC138839
34	67	3.8	147840	2	AC138939
35	67	3.8	151555	2	AC139276
36	67	3.8	151635	9	AC093308
37	67	3.8	152445	9	AC138410
38	67	3.8	152782	2	AL133255
39	67	3.8	159184	2	AC139484
40	67	3.8	161460	9	AC093243
41	67	3.8	161460	9	AL157879
42	67	3.8	161893	9	AC137781
43	67	3.8	165298	2	AC140143
44	67	3.8	168547	2	AC138947
45	67	3.8	169056	2	AC145104

ALIGNMENTS

```
RESULT 1
LOCUS      BD240720          1743 bp    DNA    linear    PAT 17-JUL-2003
DEFINITION 14273 receptor, a G-protein coupled receptor.
ACCESSION  BD240720
VERSION    BD240720.1 GI:33050490
KEYWORDS   JP 2002522011-A/1.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1743)
AUTHORS   Glucksmann, M.A. and Tsai, F.Y.
TITLE     14273 receptor, a G-protein coupled receptor
JOURNAL   Patent: JP 2002522011-A 1 23-JUL-2002;
```

Pred. No. is the number of results predicted by chance to have a

|||||
Db 1501 CTTGGAGGCTGAACTGAGAACTCTTGAACTGGGAGGCAAGGTTGAGAGCGG 1560
Qy 1561 AGATCGGCTTTCCTGCTCCCAACGAGGCAAGAGTGAATCTCTTAAAAAAA 1620
Db 1561 AGATCGGCTTTCCTGCTCCCAACGAGGCAAGAGTGAATCTCTTAAAAAAA 1620
Qy 1621 AAAAAAAGATTGTGTAGGGTCTTTTAAATGGAAGCTTTTGTAGTGTGTATA 1680
Db 1621 AAAAAAAGATTGTGTAGGGTCTTTTAAATGGAAGCTTTTGTAGTGTGTATA 1680
Qy 1681 TGATCAATTTAAATTAATTAATTTATGACTGTTCAGCAAAAAAAGG 1740
Db 1681 TGATCAATTTAAATTAATTTATTTATGACTGTTCAGCAAAAAAAGG 1740
Qy 1741 CGG 1743
Db 1741 CGG 1743
RESULT 3
AR28216 1743 bp DNA linear PAT 20-DEC-2002
LOCUS AR28216 Sequence 2 from patent US 6448005.
DEFINITION AR28216
ACCESSION AR28216
VERSION AR28216.1 GI:2726963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1743)
AUTHORS Gluckmann,M.A. and Tsai,F.-Y.
TITLE 14723 Receptor, a novel G-protein coupled receptor
JOURNAL Patent: US 6448005-A 2 10-SEP-2002;
FEATURES
Location/Qualifiers
1..1743
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1743; DB 6; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCCGGAATGTTCTAGACCGCTGGGCGCGCAGGCGCGGGAATGTCCTGAAAGCG 60
Db 1 TCCGGAATGTTCTAGACCGCTGGGCGCGCAGGCGCGGGAATGTCCTGAAAGCG 60
Qy 61 GCGGGCAGCGGGCGACGCGCTTGGCGAGCCTGAGCAAGCAACGCGCTTTCC 120
Db 61 GCGGGCAGCGGGCGACGCGCTTGGCGAGCCTGAGCAAGCAACGCGCTTTCC 120
Qy 121 CTTTCTTTCGCAAGTCAAGGGCGACACCGGCTGTGCTGGCGCGGTGAGACACCT 180
Db 121 CTTTCTTTCGCAAGTCAAGGGCGACACCGGCTGTGCTGGCGCGGTGAGACACCT 180
Qy 181 GCTGTGCTCATCTTTGCAAGTGTGCTGGGCAAGTGTGGGCGCGGTGTGTGGC 240
Db 181 GCTGTGCTCATCTTTGCAAGTGTGCTGGGCAAGTGTGGGCGCGGTGTGTGGC 240
Qy 241 GCGCGCAGACGCGCGCGCGAGCTGCTGCTGTGTAACCTCTTTCGCGGACCT 300
Db 241 GCGCGCAGACGCGCGCGCGAGCTGCTGCTGTGTAACCTCTTTCGCGGACCT 300
Qy 301 GCTCTTTCATGAGCGCTATTCCTCTGTGTGTGGCGGTGGCTGAGAGGCTGTGCT 360
Db 301 GCTCTTTCATGAGCGCTATTCCTCTGTGTGTGGCGGTGGCTGAGAGGCTGTGCT 360
Qy 361 GGGGCGCGGTGGCGCGACGCTCTTCTAGTGTAGTACCCTGAGCGCGACGCTACAT 420
Db 361 GGGGCGCGGTGGCGCGACGCTCTTCTAGTGTAGTACCCTGAGCGCGACGCTACAT 420
Qy 421 CTTGACGCTGGCCGCGGTGAGCCTGAGCGCATGTGTGACCTGTGACCTTGACCGCG 480

|||||
Db 421 CTTGACGCTGGCCGCGGTGAGCCTGAGCGCATGTGTGACCTGTGACCTTGACCGCG 480
Qy 481 GGTGCGGGGTCTGTGGGCGGGCGGGGCAAGTGTCTGTGGCGCTATCTGTGGCTATTC 540
Db 481 GGTGCGGGGTCTGTGGGCGGGCGGGGCAAGTGTCTGTGGCGCTATCTGTGGCTATTC 540
Qy 541 GGGGGTGGCGGCTGTGGCTCTGTGGCTCTTCTTGTGAGTGTGTGGCGGAGGCTCCCGG 600
Db 541 GGGGGTGGCGGCTGTGGCTCTGTGGCTCTTCTTGTGAGTGTGTGGCGGAGGCTCCCGG 600
Qy 601 GCGCGACAGAAATTTGATTTGACACATGATTTGGCCACCATTTCTGAGAGATCTC 660
Db 601 GCGCGACAGAAATTTGATTTGACACATGATTTGGCCACCATTTCTGAGAGATCTC 660
Qy 661 GTGGAGTGTCTTTTGTACTTGTGACATCTTGTGTGGCGAGCTGTGATGTGATCAG 720
Db 661 GTGGAGTGTCTTTTGTACTTGTGACATCTTGTGTGGCGAGCTGTGATGTGATCAG 720
Qy 721 TTACTCCAAATTTTACAGATCACAAGGCAATCAAGAGGCTCAGGTAAGCTGGC 780
Db 721 TTACTCCAAATTTTACAGATCACAAGGCAATCAAGAGGCTCAGGTAAGCTGGC 780
Qy 781 CTACTGAGAGCCACCAAGATCCGCGTGTCCAGACAGACTTCGCGCTTTCCGACCT 840
Db 781 CTACTGAGAGCCACCAAGATCCGCGTGTCCAGACAGACTTCGCGCTTTCCGACCT 840
Qy 841 CTTCTCTCATAGTCTCTCTTTCATCATGTGTGAGGCGCCCATCATCATCATCTTCCT 900
Db 841 CTTCTCTCATAGTCTCTCTTTCATCATGTGTGAGGCGCCCATCATCATCATCTTCCT 900
Qy 901 CATCTGATCAGAACTTCAAGCAAGACCTGTGATCTGTGCGCTCTTCTTCTTGGGT 960
Db 901 CATCTGATCAGAACTTCAAGCAAGACCTGTGATCTGTGCGCTCTTCTTCTTGGGT 960
Qy 961 GTGTGCTTGCATTTGCTAATTCAGCCCTAAACCCCATCTCTACAACTGACCTGTG 1020
Db 961 GTGTGCTTGCATTTGCTAATTCAGCCCTAAACCCCATCTCTACAACTGACCTGTG 1020
Qy 1021 CAGGAATGATGGAAGAAATTTTGTGTCTCTGTGTTCCGAAAGGAGGACATTTT 1080
Db 1021 CAGGAATGATGGAAGAAATTTTGTGTCTCTGTGTTCCGAAAGGAGGACATTTT 1080
Qy 1081 AACGACACATCTGTCAAAAGAAATGACTGTGATTAATTTGCTTAATTTTCTTAT 1140
Db 1081 AACGACACATCTGTCAAAAGAAATGACTGTGATTAATTTGCTTAATTTTCTTAT 1140
Qy 1141 AGCGAGTTTCTCACACCTGGCGAGCTGTGGCATGCTTTTAAACAGAGTTCAATTCAGT 1200
Db 1141 AGCGAGTTTCTCACACCTGGCGAGCTGTGGCATGCTTTTAAACAGAGTTCAATTCAGT 1200
Qy 1201 ACCCTCATAGTACACCTCTTAAAGAAATGAACCTATGCAAAATGACATCCAGC 1260
Db 1201 ACCCTCATAGTACACCTCTTAAAGAAATGAACCTATGCAAAATGACATCCAGC 1260
Qy 1261 GTGGTAAATTAAGGGGTGATCAACAAGTTTCAATAATTTTCCCTTATAAAGATTT 1320
Db 1261 GTGGTAAATTAAGGGGTGATCAACAAGTTTCAATAATTTTCCCTTATAAAGATTT 1320
Qy 1321 GTTGGCAGGTGAGTGTTCATGCTGTAAATCCAGACATTTTGGAGGCTGAGTGGGT 1380
Db 1321 GTTGGCAGGTGAGTGTTCATGCTGTAAATCCAGACATTTTGGAGGCTGAGTGGGT 1380
Qy 1381 GGAATCACTGAGGTGAGAGTGTGAACCAACCTGACCAATGTGTGAGACCCCGTCTC 1440
Db 1381 GGAATCACTGAGGTGAGAGTGTGAACCAACCTGACCAATGTGTGAGACCCCGTCTC 1440
Qy 1441 TACTAAATAATTAATAAATTTAGCTGGAGTGTGTGGGCACTGTATCTTACTA 1500
Db 1441 TACTAAATAATTAATAAATTTAGCTGGAGTGTGTGGGCACTGTATCTTACTA 1500
Qy 1501 CTTGGAGCTGAAACAGAGAAATCTTTGAACCTGTGGAGGCAAGGTTTCACTGACCG 1560


```

Db      1501 CTTGGAGGCTGACGAGAAATCTTGAACCTGGGAGGCAAGGTTGCAGTGAACCG 1560
Qy      1561 AGATCGCTATGCTACTCCAAACGAGGCAACAAAGTGAATCTCACTTTAAAAAAA 1620
Db      1561 AGATCGCTATGCTACTCCAAACGAGGCAACAAAGTGAATCTCACTTTAAAAAAA 1620
Qy      1621 AAAAAAAGATTGTTATATGAGGTTCTTTAAATGTGAACCTTTTACTGTTGTATATA 1680
Db      1621 AAAAAAAGATTGTTATATGAGGTTCTTTAAATGTGAACCTTTTACTGTTGTATATA 1680
Qy      1681 TGATCAAAATTTAATAATATTTATTTATGACTGTTGACGAAAAAAGG 1740
Db      1681 TGATCAAAATTTAATAATATTTATTTATGACTGTTGACGAAAAAAGG 1740
Qy      1741 CGG 1743
Db      1741 CGG 1743

RESULT 4
AR372101
LOCUS      AR372101      1743 bp      DNA      linear      PAT 12-SEP-2003
DEFINITION Sequence 2 from patent US 6395877.
ACCESSION  AR372101
VERSION     AR372101.1 GI:34609378
KEYWORDS
SOURCE      Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1743)
            Glucksmann,M.A. and Tsai,P.-Y.
            14273 receptor, a novel G-protein coupled receptor
            Patent: US 6395877-A 2 28-MAY-2002;
            Location/Qualifiers
FEATURES
            source
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 1743; DB 6; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCGGAGCTAGTTCTTAAGCCGCTGCGGCGCCGACGCGCGGAAATGCTCCCTGAATGCG 60
Db      1 TCGGAGCTAGTTCTTAAGCCGCTGCGGCGCCGACGCGCGGAAATGCTCCCTGAATGCG 60
Qy      61 GCGGAGGAGCGGAGCGCGGCTTTGGGCAAGCTGAGCAAGCAACCGGCTTTC 120
Db      61 GCGGAGGAGCGGAGCGCGGCTTTGGGCAAGCTGAGCAAGCAACCGGCTTTC 120
Qy      121 CTTCTTCTCGAGCTGAAGGCGCAACCGGCTGCTGCGCGGCTGAGCAACCGT 180
Db      121 CTTCTTCTCGAGCTGAAGGCGCAACCGGCTGCTGCGCGGCTGAGCAACCGT 180
Qy      181 GCTGCTGCTATTTTTCAGTGTGCTGCTGCGGCAAGCTGCTGCTGCTGCTGCTGCT 240
Db      181 GCTGCTGCTATTTTTCAGTGTGCTGCTGCGGCAAGCTGCTGCTGCTGCTGCTGCT 240
Qy      241 GCGGCGAGCGAGCGGCGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db      241 GCGGCGAGCGAGCGGCGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy      301 GCTCTTCAATAGCGCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db      301 GCTCTTCAATAGCGCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy      361 GGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db      361 GGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy      421 CCTCAGCTGAGCGGCTGAGCGCTGAGCGCATGCTGATGCTGACCTGCAAGCGCG 480
Db      421 CCTCAGCTGAGCGGCTGAGCGCTGAGCGCATGCTGATGCTGACCTGCAAGCGCG 480

Db      421 CTTCAAGCTGAGCGGCTGAGCGCTGAGCGCATGCTGATGCTGACCTGCAAGCGCG 480
Qy      481 GGTGGGGGTCTGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db      481 GGTGGGGGTCTGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Qy      541 GGGGGGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db      541 GGGGGGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy      601 CGCGGCAAGAGAAATTTGATTTGCAACATGATTTGGCCCAACATTTCTGAGAGATCTC 660
Db      601 CGCGGCAAGAGAAATTTGATTTGCAACATGATTTGGCCCAACATTTCTGAGAGATCTC 660
Qy      661 GTGGGATGCTCTTTTGTATCTTTGAACTTTGAGCTTGGTGCAGAGCTGCTCATTTGATCAG 720
Db      661 GTGGGATGCTCTTTTGTATCTTTGAACTTTGAGCTTGGTGCAGAGCTGCTCATTTGATCAG 720
Qy      721 TTAATCCAAAATTTTACAGATCAAGAGGATCAAGGAGAGAGGCTACGGTTAAGCCTGGC 780
Db      721 TTAATCCAAAATTTTACAGATCAAGAGGATCAAGGAGAGAGGCTACGGTTAAGCCTGGC 780
Qy      781 CTACTGAGAGGCAACAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db      781 CTACTGAGAGGCAACAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy      841 CTTCTCTCAGTGTCTCTTCTTCAATCATGTGAGGCCCATCATCATCTCTCT 900
Db      841 CTTCTCTCAGTGTCTCTTCTTCAATCATGTGAGGCCCATCATCATCTCTCT 900
Qy      901 CATCCGATCCAGAACTTCAAGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db      901 CATCCGATCCAGAACTTCAAGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy      961 GGTGGCTTCACTTGTGCTAATCAGCCCTTAACCCCATCTCTTACAACTGACACTGTG 1020
Db      961 GGTGGCTTCACTTGTGCTAATCAGCCCTTAACCCCATCTCTTACAACTGACACTGTG 1020
Qy      1021 CAGGAATGAGTGAAGAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db      1021 CAGGAATGAGTGAAGAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy      1081 AACAGACACTCTGTAAAGAAATGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db      1081 AACAGACACTCTGTAAAGAAATGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy      1141 AGCGAGTTTCTCAACTGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db      1141 AGCGAGTTTCTCAACTGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy      1201 ACCCTCATAGTGCAACCTGCTTTAAGAAATGAACCTATGCAATGACATCCACAGC 1260
Db      1201 ACCCTCATAGTGCAACCTGCTTTAAGAAATGAACCTATGCAATGACATCCACAGC 1260
Qy      1261 GTGGGTAAATTAAGGGGTATCAACAAAGTTCAATAATTTTCCCTTAAAGAAATTT 1320
Db      1261 GTGGGTAAATTAAGGGGTATCAACAAAGTTCAATAATTTTCCCTTAAAGAAATTT 1320
Qy      1321 GTTGGCAGGTGAGGTGCTCATGCTGTAATCCAGAGTTTGGAGGCTGAGGTGAGT 1380
Db      1321 GTTGGCAGGTGAGGTGCTCATGCTGTAATCCAGAGTTTGGAGGCTGAGGTGAGT 1380
Qy      1381 GATTCACCTGAGGTGAGGTGCTGAGACCACTGACCAAGTGTGAGACCTCCGCTC 1440
Db      1381 GATTCACCTGAGGTGAGGTGCTGAGACCACTGACCAAGTGTGAGACCTCCGCTC 1440
Qy      1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
Db      1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
Qy      1501 CTTGGAGGCTGAACCAAGAGAAATCTTTGAACCTGAGGCGAGAGTTGCACTGAGCG 1560
Db      1501 CTTGGAGGCTGAACCAAGAGAAATCTTTGAACCTGAGGCGAGAGTTGCACTGAGCG 1560
```

QY 1561 AGATCGTCATTGCACTCCACACGAGGCAACAAGAGTGAACCTCATCTTAAAAAAA 1620
Db 1561 AGATCGTCATTGCACTCCACACGAGGCAACAAGAGTGAACCTCATCTTAAAAAAA 1620
QY 1621 AAAAAAAGATTGTAAGGCTTCTTTAAATGTGAACCTTTTAACTGTGTGTAATA 1680
Db 1621 AAAAAAAGATTGTAAGGCTTCTTTAAATGTGAACCTTTTAACTGTGTGTAATA 1680
QY 1681 TGAATCAATTTAATTAATATTATTATTATTTGACTGTTCAGCAAAAAAAGGG 1740
Db 1681 TGAATCAATTTAATTAATATTATTATTATTTGACTGTTCAGCAAAAAAAGGG 1740
QY 1741 CGG 1743
Db 1741 CGG 1743

RESULT 5
BD015337
LOCUS BD015337 1458 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel polypeptide.
ACCESSION BD015337 GI:22556475
VERSION BD015337.1 JP 2001211885-A/1.
KEYWORDS JP 2001211885-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Yoshimatsu, H., Sasaki, K., Nakatani, Y., Saeiki, S., Miura, K. and Saeiki, S.
REFERENCE Patent: JP 2001211885-A 1 07-AUG-2001;
JOURNAL Kiyoma HAKKO KOGYO CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2001211885-A/1
PD 07-AUG-2001
PF 02-FEB-2000 JP 2000024921
PI HAJIME YOSHIMATSU, KATSUOSHI SASAKI, YUKIE NAKATANI, SATOSHI PI
SAEIKI, S.
PI KAZUMI MIURA, SUSUMU SEKINE
PC C12N15/09, A01H5/00, A01K67/027, A61K38/00, A61K39/395, A61K39/395, A61K39/395.
PC A61K45/00, A61P5/08, A61P35/00, C07K14/705, C07K16/28, C12N1/15, PC C12N1/19, C12N5/10, C12N5/10, C12P21/02, C12P21/08, C12P1/68 PC C12P1/68, G01N33/15, G01N33/50, G01N33/566, C12N15/00, A61K37/02, C12N5/00, C12N5/00 CC G01N33/50, G01N33/566, C12N15/00, A61K37/02, C12N5/00, C12N5/00 CC
FH Key Location/Qualifiers
FEATURES
source 1.1458
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 76.2%; Score 1328; DB 6; Length 1458;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 195 TTGAGTGTGCTGCTGAGGCAAGTGTGCGCCCTGCTGTGTGAGCGCCGACGAGCC 254
Db 209 TTGAGTGTGCTGCTGAGGCAAGTGTGCGCCCTGCTGTGTGAGCGCCGACGAGCC 268
QY 255 GCGGCGGACTGCTGCTGTGTAAGTGAACCTCTTTCGCGGAGCTGCTTTATACAGC 314
Db 269 GCGGCGGACTGCTGCTGTGTAAGTGAACCTCTTTCGCGGAGCTGCTTTATACAGC 328
QY 315 CTATCCCTGAGTGTGCTGAGCGGCTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 374
Db 329 CTATCCCTGAGTGTGCTGAGCGGCTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 388
QY 375 GCCACCTGCTCTTCTAAGTATGACCTTGAAGCGGAGCTGACCTATCTTCAAGCTGAGC 434
Db 389 GCCACCTGCTCTTCTAAGTATGACCTTGAAGCGGAGCTGACCTATCTTCAAGCTGAGC 448
QY 435 CGGTGAGCTGTGAGCGGAGCTGAGTGTGATGTGACCTGACGCGGCGCTGCGGAGCTGAG 494
Db 449 CGGTGAGCTGTGAGCGGAGCTGAGTGTGATGTGACCTGACGCGGCGCTGCGGAGCTGAG 508
QY 495 GCGGCGGAGCGGAGCTGAGTGTGATGTGACCTGAGCGCTATCTGAGGCTATTTGCGAGCTG 554
Db 509 GCGGCGGAGCGGAGCTGAGTGTGATGTGACCTGAGCGCTATCTGAGGCTATTTGCGAGCTG 568
QY 555 TGCCTCTGTGCTCTTCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 614
Db 569 TGCCTCTGTGCTCTTCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 628
QY 615 TTTCGATTGACACTGATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 674
Db 629 TTTCGATTGACACTGATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 688
QY 675 TTGTTCTTTGAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
Db 689 TTGTTCTTTGAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
QY 735 TACAGATCAAGAGGATCAAGAGAGGCTACGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 794
Db 749 TACAGATCAAGAGGATCAAGAGAGGCTACGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 808
QY 795 ACCAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
Db 809 ACCAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 868
QY 855 TCTCCTTCTTCAATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 914
Db 869 TCTCCTTCTTCAATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
QY 915 ACTTCAAGCAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
Db 929 ACTTCAAGCAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
QY 975 TTGCTAATTAAGGCTTAAACCTCAATCTCTAACAATGACATGCTGAGGAATGAGTGA 1034
Db 989 TTGCTAATTAAGGCTTAAACCTCAATCTCTAACAATGACATGCTGAGGAATGAGTGA 1048
QY 1035 AGAAAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1094
Db 1049 AGAAAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1108
QY 1095 TCAAAAAATTAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1154
Db 1109 TCAAAAAATTAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1168
QY 1155 CACTGCGAGCTGTGCTGCTGCTTAAACAGAGTTTCAATTTCCAGTACCTTCATCAGTG 1214
Db 1169 CACTGCGAGCTGTGCTGCTGCTTAAACAGAGTTTCAATTTCCAGTACCTTCATCAGTG 1228
QY 1215 CACCTGCTTTAAGAAATGAACCTTATGCAATATGACATCAGAGCTGCTGCTGCTGCTG 1274
Db 1229 CACCTGCTTTAAGAAATGAACCTTATGCAATATGACATCAGAGCTGCTGCTGCTGCTG 1288
QY 1275 GGGTGATCAACAAATTTTCAATATATTTTCCCTTATTAAGAAAGATTGTTGGCAGGTGCA 1334

Db 1289 GGGTATACCAAGTTTCAATATTTTCCCTTATTAATAAGATTGTTGGCAGAGTCA 1348
 Qy 1335 GTGGTTCATGCTCTGTAATCCAGAGTTTGGAGGCTGAGTGGGTGATCACTGAGGT 1394
 Db 1349 GTGGTTCATGCTCTGTAATCCAGAGTTTGGAGGCTGAGTGGGTGATCACTGAGGT 1408
 Qy 1395 CAGAGTTTGAGACCAACCTGACCAATGTTGAGACCCCGCTCTACT 1444
 Db 1409 CAGAGTTTGAGACCAACCTGACCAATGTTGAGACCCCGCTCTACT 1458

RESULT 6
 AX168141
 LOCUS AX168141 1321 bp DNA linear PAT 03-JUL-2001
 DEFINITION Sequence 59 from Patent WO0142288.
 AX168141
 VERSION AX168141.1 GI:14597419
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 1 Burford, N., Baughn, M.R., Au-Young, J., Yang, J., Lu, D.A. and Reddy, R.
 G-protein coupled receptors
 Patent: WO 0142288-A 59 14-JUN-2001;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 1. 1321
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 5029478CB1"

ORIGIN

Query Match 62.0%; Score 1081; DB 6; Length 1321;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1181; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 32 CAGGCCCGGAAATGTCCTGATGCGCGGAGCAGCGGCGACGCGCTTGGCAGC 91
 Db 18 CAGGCCCGGAAATGTCCTGATGCGCGGAGCAGCGGCGACGCGCTTGGCAGC 77
 Qy 92 CTGGAGCAAGCCACCGACCGCTTCCCTTCTTCTCGAGCTCAAGGCGACCAACCG 151
 Db 78 CTGGAGCAAGCCACCGACCGCTTCCCTTCTTCTCGAGCTCAAGGCGACCAACCG 137
 Qy 152 CTGGTCTGGCGCGGTGAGACACCGCTGCTGCTCACTTTCAGATGCTGCTG 211
 Db 138 CTGGTCTGGCGCGGTGAGACACCGCTGCTGCTCACTTTCAGATGCTGCTG 197
 Qy 212 GGCACAGTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271
 Db 198 GGCACAGTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257
 Qy 272 CTGGTCTGCAACTCTTCTGCGGACCTGCTTCTATCAGCGCTATCCTCTGCTG 331
 Db 258 CTGGTCTGCAACTCTTCTGCGGACCTGCTTCTATCAGCGCTATCCTCTGCTG 317
 Qy 332 GCGGTGCGGTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
 Db 318 GCGGTGCGGTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
 Qy 392 GTGATGACCTGAGCGGACGCTCAGCATCTCAGCGTGGCGCGGCTGAGCTGAGCG 451
 Db 378 GTGATGACCTGAGCGGACGCTCAGCATCTCAGCGTGGCGCGGCTGAGCTGAGCG 437
 Qy 452 ATGGTGTGATCTGTGCACTGCGACCGGCGGTGCGGGGCTTCTGAGCGCGCGGCA 511
 Db 438 ATGGTGTGATCTGTGCACTGCGACCGGCGGTGCGGGGCTTCTGAGCGCGCGGCA 497
 Qy 512 GTGCTGTGCGCGCTCATCTGGGCGGTATGCGGCGGTGCGCGCTGCTGCTGCTGCT 571

Db 498 GTGCTGTGCGCGCTCATCTGGGCGGTATGCGGCGGTGCGCGCTGCTGCTGCTGCT 557
 Qy 572 TTTCAGTGTGCTCCGCAACGCTCCCGGCGCGGACCAAGAAATTTGATTTGACACTG 631
 Db 558 TTCCAGTGTGCTCCGCAACGCTCCCGGCGCGGACCAAGAAATTTGATTTGACACTG 617
 Qy 632 ATTTGGCCCAACCTTCTGAGAGATCTGTTGGAGTCTCTTTGTTACTTTGAATTC 691
 Db 618 ATTTGGCCCAACCTTCTGAGAGATCTGTTGGAGTCTCTTTGTTACTTTGAATTC 677
 Qy 692 TTGTCACAGAGATCTGATCTGATCACTTCTCAAAATTTTACAGATCAAGAGCA 751
 Db 678 TTGTCACAGAGATCTGATCTGATCACTTCTCAAAATTTTACAGATCAAGAGCA 737
 Qy 752 TCAAGAGAGGCTCAGCGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
 Db 738 TCAAGAGAGGCTCAGCGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
 Qy 812 CAGCAGAGCTTCCGCTTCCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871
 Db 798 CAGCAGAGCTTCCGCTTCCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
 Qy 872 TGGAGCCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 931
 Db 858 TGGAGCCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 917
 Qy 932 GTCATCTGCGCGCT 991
 Db 918 GTCATCTGCGCGCT 977
 Qy 992 AACCCCATCTCTCAACATGACATGACATGACATGACATGACATGACATGACATGAC 1051
 Db 978 AACCCCATCTCTCAACATGACATGACATGACATGACATGACATGACATGACATGAC 1037
 Qy 1052 TTCTGTTCCAGAAAGGAGGCTTTTAAACAGACATCTGTCAAAAGAAATGACTTG 1111
 Db 1038 TTCTGTTCCAGAAAGGAGGCTTTTAAACAGACATCTGTCAAAAGAAATGACTTG 1097
 Qy 1112 TCGATTATTTCTGGCTAATTTTCTTATAGCGGAGTTTCTCAACCTGCGAGCTGTG 1171
 Db 1098 TCGATTATTTCTGGCTAATTTTCTTATAGCGAGTTTCTCAACCTGCGAGCTGTG 1157
 Qy 1172 CATGCTTTTAAAGAGTTTCAATTTCCAGTACCTTCATAGTG 1214
 Db 1158 CATGCTTTTAAAGAGTTTCAATTTCCAGTACCTTCATAGTG 1200

RESULT 7
 AL356214
 LOCUS AL356214 163964 bp DNA linear PRI 13-APR-2001
 DEFINITION Human DNA sequence from clone RP11-30816 on chromosome 10, complete
 sequence.
 AL356214
 VERSION AL356214.20 GI:13624997
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 1 (bases 1 to 163964)
 Brown, A.
 Direct Submision
 Submitted (13-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Apr 14, 2001 this sequence version replaced gi:13160256.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWSRPT; Tr, TRMBL; Wp, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10> RP11-30E16 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

This sequence is the entire insert of clone RP11-30E16. The true left end of clone RP11-43J02 is at 129920 in this sequence. The true right end of clone RP11-313N18 is at 82571 in this sequence.

FEATURES
Location/Qualifiers

1..163964

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="10"

/clone_id="RP11-30E16"

/clone_lib="RP11-11.1"

repeat_region repeat: matches 5968..6084 of consensus"

repeat_region repeat: matches 1..299 of consensus"

repeat_region repeat: matches 5793..5946 of consensus"

repeat_region repeat: matches 1..301 of consensus"

repeat_region repeat: matches 5417..5793 of consensus"

repeat_region repeat: matches 3..302 of consensus"

repeat_region repeat: matches 8..140 of consensus"

repeat_region repeat: matches 5232..5552 of consensus"

repeat_region repeat: matches 1183..1645 of consensus"

repeat_region repeat: matches 1..312 of consensus"

repeat_region repeat: matches 1097..1183 of consensus"

repeat_region repeat: matches 695..969 of consensus"

repeat_region repeat: matches 18..283 of consensus"

repeat_region repeat: matches 4..92 of consensus"

repeat_region repeat: matches 2458..2704 of consensus"

repeat_region repeat: matches 2648..2708 of consensus"
/note="L2 repeat: matches 20..252 of consensus"
repeat_region repeat: matches 1..133 of consensus"
/note="FRAM_C repeat: matches 2544..2698 of consensus"
/note="L2 repeat: matches 2317..2412 of consensus"
repeat_region repeat: matches 9424..9511
/note="L2 repeat: matches 2210..2585 of consensus"
repeat_region repeat: matches 9752..10112
/note="Charles repeat: matches 2210..2585 of consensus"
repeat_region repeat: matches 10113..10154
/note="21 copies 2 mer aa 78% conserved"
repeat_region repeat: matches 10173..10333
/note="MERS repeat: matches 3..154 of consensus"
repeat_region repeat: matches 11413..11528
/note="FRAM_C repeat: matches 1..116 of consensus"
repeat_region repeat: matches 11536..11719
/note="MIR repeat: matches 1..192 of consensus"
repeat_region repeat: matches 12017..12295
/note="Alu repeat: matches 25..276 of consensus"
repeat_region repeat: matches 12527..13073
/note="CpG island"
/evidence="not_experimental"
repeat_region repeat: matches 13079..13130
/note="MIR repeat: matches 87..151 of consensus"
repeat_region repeat: matches 13409..13715
/note="AluX repeat: matches 1..307 of consensus"
repeat_region repeat: matches 13987..14269
/note="AluX repeat: matches 1..280 of consensus"
repeat_region repeat: matches 14293..14330
/note="19 copies 2 mer aa 81% conserved"
repeat_region repeat: matches 14837..15042
/note="L2 repeat: matches 2541..2743 of consensus"
repeat_region repeat: matches 15659..15969
/note="AluX repeat: matches 1..301 of consensus"
repeat_region repeat: matches 16912..17073
/note="FRAM repeat: matches 1..161 of consensus"
repeat_region repeat: matches 17557..17845
/note="AluX repeat: matches 1..310 of consensus"
repeat_region repeat: matches 20154..20274
/note="MIR repeat: matches 118..259 of consensus"
repeat_region repeat: matches 20341..20635
/note="AluX repeat: matches 1..295 of consensus"
repeat_region repeat: matches 20960..21261
/note="AluX repeat: matches 1..305 of consensus"
repeat_region repeat: matches 21583..21657
/note="MIR repeat: matches 113..200 of consensus"
repeat_region repeat: matches 22002..22045
/note="Alu repeat: matches 248..291 of consensus"
repeat_region repeat: matches 22046..22168
/note="AluX repeat: matches 14..136 of consensus"
repeat_region repeat: matches 22170..22391
/note="AluX repeat: matches 1..235 of consensus"
repeat_region repeat: matches 22717..23023
/note="AluX repeat: matches 1..301 of consensus"
repeat_region repeat: matches 23376..23670
/note="AluX repeat: matches 2..301 of consensus"
repeat_region repeat: matches 23786..24096
/note="AluX repeat: matches 1..308 of consensus"
repeat_region repeat: matches 24754..25151
/note="L2 repeat: matches 2290..2700 of consensus"
repeat_region repeat: matches 25514..25615
/note="MERS repeat: matches 2..105 of consensus"
repeat_region repeat: matches 25619..25648
/note="10 copies 3 mer att 100% conserved"
repeat_region repeat: matches 25650..25935
/note="AluX repeat: matches 1..286 of consensus"
repeat_region repeat: matches 26484..26738
/note="LINC4 repeat: matches 7762..7974 of consensus"
repeat_region repeat: matches 26741..26830
/note="S repeat: matches 1..90 of consensus"

```

repeat_region      26831..26964
                    /note="LMC4 repeat: matches 7641..7775 of consensus"
repeat_region      27088..27386
                    /note="AluSg repeat: matches 3..299 of consensus"
repeat_region      27741..28051
                    /note="AluSg repeat: matches 1..300 of consensus"
repeat_region      28061..28094
                    /note="17 copies 2 mer aa 82% conserved"
repeat_region      28102..28390
                    /note="AluSx repeat: matches 6..296 of consensus"
repeat_region      28393..28687
                    /note="AluSg repeat: matches 1..306 of consensus"
repeat_region      29954..30100
                    /note="AluSg repeat: matches 82..231 of consensus"
repeat_region      30101..30380
                    /note="AluSx repeat: matches 1..294 of consensus"
repeat_region      30381..30458
                    /note="MLTIC repeat: matches 6..83 of consensus"
repeat_region      30573..30709
                    /note="LRR16B repeat: matches 319..462 of consensus"
repeat_region      30919..31222
                    /note="AluSg repeat: matches 1..305 of consensus"
repeat_region      31234..31279
                    /note="LRR16A repeat: matches 1..45 of consensus"
repeat_region      32166..32272
                    /note="MIR repeat: matches 124..228 of consensus"
repeat_region      32606..32639
                    /note="17 copies 2 mer tc 88% conserved"
repeat_region      33420..33763
                    /note="L2 repeat: matches 2325..2742 of consensus"
repeat_region      34296..34599

```

```

Query Match      56.6%; Score 986; DB 9; Length 163964;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 986; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 ACAGATCACAAGGCTCAAGAGGCTCAGGTAAAGCTGCGCTACTCGAGAGCCA 795
DB 103366 ACAGATCACAAGGCTCAAGAGGCTCAGGTAAAGCTGCGCTACTCGAGAGCCA 103425
QY 796 CCAGATCCGCGTCCAGAGAGACTTCCGCTCTTCCGAGCCCTCTTCCCTCAATGT 855
DB 103426 CCAGATCCGCGTCCAGAGAGACTTCCGCTCTTCCGAGCCCTCTTCCCTCAATGT 103485
QY 856 CTCCTTCTTATCATGTAGAGAGCCCATCATCATCACCATCTCTCTCAATCTGATCCAGAA 915
DB 103486 CTCCTTCTTATCATGTAGAGAGCCCATCATCATCACCATCTCTCTCAATCTGATCCAGAA 103545
QY 916 CTTCAAGCAAGACTGTGATCTGCGCGCTCTCTTCTTGGGTGTGGCTTTCACATT 975
DB 103546 CTTCAAGCAAGACTGTGATCTGCGCGCTCTCTTCTTGGGTGTGGCTTTCACATT 103605
QY 976 TGTCTAATTCAGCCCTAAACCCCATCTCTCAACATGACACTGTGACGGAATGATGGA 1035
DB 103606 TGTCTAATTCAGCCCTAAACCCCATCTCTCAACATGACACTGTGACGGAATGATGGA 103665
QY 1036 GAAAAATTTTGTGCTTGTGTTCCAGAAAAAGGAGCCATTTTAAACAGACATCTGT 1095
DB 103666 GAAAAATTTTGTGCTTGTGTTCCAGAAAAAGGAGCCATTTTAAACAGACATCTGT 103725
QY 1096 CAAAAAATGATTTGCTGATTTATTTCTGGCTAATTTTCTTATAGCCAGTTTCTCAC 1155
DB 103726 CAAAAAATGATTTGCTGATTTATTTCTGGCTAATTTTCTTATAGCCAGTTTCTCAC 103785
QY 1156 ACCTGGAGAGCTGTGAGATCTTTTAAACAGAGTCAATTTCCAGTACCTCCATCAGTGC 1215
DB 103786 ACCTGGAGAGCTGTGAGATCTTTTAAACAGAGTCAATTTCCAGTACCTCCATCAGTGC 103845
QY 1216 ACCCTGCTTTAAGAAAAATGAACCTATGCAAAATGACATCCAGAGCTGCGTAAATTAAGG 1275
DB 103846 ACCCTGCTTTAAGAAAAATGAACCTATGCAAAATGACATCCAGAGCTGCGTAAATTAAGG 103905
QY 1276 GGTGATCACCAGTTTATATATTTTCCCTTATATAAGGATTTGTGGCCAGGTGCAG 1335

```

```

DB 103906 GGTGATCACCAGTTTATATATTTTCCCTTATATAAGGATTTGTGGCCAGGTGCAG 103965
QY 1336 TGTTCATGCTGCTTAATCCAGCAGATTGGAGAGCTGAGGTGGTGATCACTGAGTTC 1395
DB 103966 TGTTCATGCTGCTTAATCCAGCAGATTGGAGAGCTGAGGTGGTGATCACTGAGTTC 104025
QY 1396 AGAGATTCAGACCAACCTGACCAATGATGTGAGACCCCGCTCTAATAAATAA 1455
DB 104026 AGAGATTCAGACCAACCTGACCAATGATGTGAGACCCCGCTCTAATAAATAA 104085
QY 1456 AAAAAATAGCTGGAGTGTGTTGGGACCTGTAATCTCAGTACTTGGAGAGCTGAC 1515
DB 104086 AAAAAATAGCTGGAGTGTGTTGGGACCTGTAATCTCAGTACTTGGAGAGCTGAC 104145
QY 1516 CAGAGAAATCTTGAACCTGGAGGACAGAGTTGACAGTGGAGCCGAGATCTGCCATTGC 1575
DB 104146 CAGAGAAATCTTGAACCTGGAGGACAGAGTTGACAGTGGAGCCGAGATCTGCCATTGC 104205
QY 1576 ACTCCAAACGAGGCAACAGAGTGAATCTCACTTTAAAAAAGATTTGT 1635
DB 104206 ACTCCAAACGAGGCAACAGAGTGAATCTCACTTTAAAAAAGATTTGT 104265
QY 1636 TATGGTCTCTTAAATGAGACTTTTGTGTTGTGTTGTAATATGATCAAAATTTATA 1695
DB 104266 TATGGTCTCTTAAATGAGACTTTTGTGTTGTGTTGTAATATGATCAAAATTTATA 104325
QY 1696 AATATTTATTTATGACTGTTCAGCAA 1721
DB 104326 AATATTTATTTATGACTGTTCAGCAA 104351

```

```

RESULT 8
AX498184 1086 bp DNA linear PAT 26-SEP-2002
LOCUS AX498184
DEFINITION Sequence 7 from Patent WO242461.
ACCESSION AX498184
VERSION AX498184.1 GI:23343113
KEYWORDS
ORGANISM
SOURCE
REFERENCE
1. Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
Endogenous and non-endogenous versions of human g protein-coupled
receptors
Patent: WO 0242461-A 7 30-MAY-2002;
Arena Pharmaceuticals, Inc. (US)
FEATURES
Source
1..1086
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Novel Sequence"
ORIGIN

```

```

Query Match      53.1%; Score 926; DB 6; Length 1086;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 51 CTGAATGCGCGCGGAGAGCGGCGGAGCGGCGCTTGGCGAGCTTGGAGCAAGCCGCA 110
DB 8 CTGAATGCGCGCGGAGAGCGGCGGAGCGGCGCTTGGCGAGCTTGGAGCAAGCCGCA 67
QY 111 CCGCTTTCCTTCTCTCCAGCTCAAGGAGGAGCAGCAGGCTGTGCTGCGCGGCTGG 170
DB 68 CCGCTTTCCTTCTCTCCAGCTCAAGGAGGAGCAGCAGGCTGTGCTGCGCGGCTGG 127
QY 171 AGACAAACCGTGTGCTGCTATCTTTTGAAGTGTGCTGCTGAGCAAGTGTGCGCCCTGG 230
DB 128 AGACAAACCGTGTGCTGCTATCTTTTGAAGTGTGCTGCTGAGCAAGTGTGCGCCCTGG 187
QY 231 TGTGTGCGCGCGGAGCAGACCGCGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCT 290

```



```

QY 788 GAGAGCCACAGATCCGCGTGTCCAGCAGGACTTCCGCGCTCTTCCGACCTCTTCTC 847
DB 601 GAGAGCCACAGATCCGCGTGTCCAGCAGGACTTCCGCGCTCTTCCGACCTCTTCTC 660
QY 848 CTGATGCTCTCTTCTTATCATGTGAGAGCCCATCATCATCATCATCTCTCTCATCTG 907
DB 661 CTGATGCTCTCTTCTTATCATGTGAGAGCCCATCATCATCATCATCTCTCTCATCTG 720
QY 908 ATCCAGAACTTCAAGCAAGACTG 931
DB 721 ATCCAGAACTTCAAGCAAGACTG 744

RESULT 10
AX647273 1066 bp DNA linear PAT 04-MAR-2003
LOCUS AX647273
DEFINITION Sequence 1465 from Patent EP1270724.
ACCESSION AX647273
VERSION AX647273.1 GI:28880736
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
TITLE Guanosine triphosphate-binding protein coupled receptors
JOURNAL Patient: EP 1270724-A 1465 02-JAN-2003;
        National Institute of Advanced Industrial Science and Technology
        (JIP) ; Center for Advanced Science and Technology Incubation, Ltd.
        (JIP)

FEATURES
    source
        1..1066
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            201..866
            /note="unassigned protein product"
            /codon_start=1
            /protein_id="CAD69757.1"
            /db_xref="GI:28880737"
            /db_xref="REFSEQ:CAD69757"
            /translation="MSPEBCARAGADAPLRLEQANRRFPFSDYKDHVLVAAYET
            TVLVLFVSLGNVCLVVARRRRGATDVLNLFCAULFTSATPLVAVRWTE
            AMVLGPVACHLFFVMTLSGVTILTLAAVSLERMCIVHLOGVRGPRARAVLVA
            LIMGSAVALPLCFVRVVPRLPGADQVSAFLCVGRCPAQAGKRGPDGSDDEDDQ
            EQG"

CDS
    1..1066
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        201..866
        /note="unassigned protein product"
        /codon_start=1
        /protein_id="CAD69757.1"
        /db_xref="GI:28880737"
        /db_xref="REFSEQ:CAD69757"
        /translation="MSPEBCARAGADAPLRLEQANRRFPFSDYKDHVLVAAYET
        TVLVLFVSLGNVCLVVARRRRGATDVLNLFCAULFTSATPLVAVRWTE
        AMVLGPVACHLFFVMTLSGVTILTLAAVSLERMCIVHLOGVRGPRARAVLVA
        LIMGSAVALPLCFVRVVPRLPGADQVSAFLCVGRCPAQAGKRGPDGSDDEDDQ
        EQG"

ORIGIN
Query Match 32.1%; Score 559; DB 6; Length 1066;
Best Local Similarity 100.0%; Pred. No. 2,7e-313;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 315 CTATCCCTCTGATGCGCGGCGTGGAGCTGAGAGGCTGAGGCGCCGCTTGCCT 374
DB 472 CTATCCCTCTGATGCGCGGCGTGGAGCTGAGAGGCTGAGGCGCCGCTTGCCT 531
QY 375 GCCACCTGCTCTTCTAAGTATGACCTGAGCGGACGCTCACCATCTCAAGCTGACG 434
DB 532 GCCACCTGCTCTTCTAAGTATGACCTGAGCGGACGCTCACCATCTCAAGCTGACG 591
QY 435 CGGTCAAGCTTGAAGCGGATGTGTGATCGTGACCTGACGCGGCGGTGCGGCTCTG 494
DB 592 CGGTCAAGCTTGAAGCGGATGTGTGATCGTGACCTGACGCGGCGGTGCGGCTCTG 651
QY 495 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 554
DB 652 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 711
QY 555 TGCTCTCTGCTGCTTCTT 573
DB 712 TGCTCTCTGCTGCTTCTT 730

RESULT 11
AY288417 1131 bp mRNA linear PRI 08-DEC-2003
LOCUS AY288417
DEFINITION Homo sapiens G protein-coupled receptor 120 (GPR120) mRNA, complete
cDNA.
ACCESSION AY288417
VERSION AY288417.1 GI:32165517
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Fredriksson,R., Hoglund,P.J., Gloriam,D.E., Lagerstrom,M.C. and
        Schioth,H.B.
TITLE Seven evolutionarily conserved human rhodopsin G protein-coupled
        receptors lacking close relatives
JOURNAL FEBS Lett. 554 (3), 381-388 (2003)
MEDLINE 14623098
PUBMED 22985413
REFERENCE
AUTHORS Fredriksson,R., Hoglund,P.J., Gloriam,D.E.I., Lagerstrom,M.C. and
        Schioth,H.B.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) Neuroscience; Unit of Pharmacology, Uppsala
        University, Box 593, Uppsala 75124, Sweden

FEATURES
    source
        1..1131
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            1..1131
            /gene="GPR120"
            1..1131
            /gene="GPR120"
            /codon_start=1
            /product="G protein-coupled receptor 120"
            /protein_id="AAP22126.1"
            /db_xref="GI:32165518"
            /translation="MSPEBCARAGADAPLRLEQANRRFPFSDYKDHVLVAAYET
            TVLVLFVSLGNVCLVVARRRRGATDVLNLFCAULFTSATPLVAVRWTE
            AMVLGPVACHLFFVMTLSGVTILTLAAVSLERMCIVHLOGVRGPRARAVLVA
            LIMGSAVALPLCFVRVVPRLPGADQVSAFLCVGRCPAQAGKRGPDGSDDEDDQ
            EQG"

ORIGIN
Query Match 25.5%; Score 445; DB 9; Length 1131;
Best Local Similarity 99.8%; Pred. No. 7.6e-247;
Matches 495; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


Qy	771	TAAAGCTGAGCCCTAACTCGAGAGACACACATACGGGTGTCACAGAGAGACTTCGGAGTCT	830
Db	771	TAAAGCTGAGCCCTAACTCGAGAGACACACATACGGGTGTCACAGAGAGACTTCGGAGTCT	830
Qy	831	TCCGACACCTCTTCCTCCTCATAGTGTCTCCTTCTTCATCATATGAGAGCCCATCATCATCA	890
Db	831	TCCGACACCTCTTCCTCCTCATAGTGTCTCCTTCTTCATCATATGAGAGCCCATCATCATCA	890
Qy	891	CCATCTCTCTCATCTCTGATCCAGAACTTCAGCAAGACCTGGTCACTTGGCCGTCTCT	950
Db	891	CCATCTCTCTCATCTCTGATCCAGAACTTCAGCAAGACCTGGTCACTTGGCCGTCTCT	950
Qy	951	TCTTCTGGGTGGTGGCCCTTCAATTGTGCTAATTCAGCCCTTAACCCCATCTCTCAACA	1011
Db	951	TCTTCTGGGTGGTGGCCCTTCAATTGTGCTAATTCAGCCCTTAACCCCATCTCTCAACA	1011
Qy	1011	TGACACTGTGACAGAAATGATGGAAGAAAAATTTTGTGCTGTCTTGTTCCAGAAAAAG	1071
Db	1011	TGACACTGTGACAGAAATGATGGAAGAAAAATTTTGTGCTGTCTTGTTCCAGAAAAAG	1071
Qy	1071	GAGCCATTTTAAACAGACAATCTGTCAAAAAGAAATGACTGTGATTAATTTCTGGCTAAT	1131
Db	1071	GAGCCATTTTAAACAGACAATCTGTCAAAAAGAAATGACTGTGATTAATTTCTGGCTAAT	1131
Qy	1131	TTTTCTTTATATAGCCGATTTCTCAACCTCGGAGACTGTGGATGCTTTTAAACAGAGTT	1191
Db	1131	TTTTCTTTATATAGCCGATTTCTCAACCTCGGAGACTGTGGATGCTTTTAAACAGAGTT	1191
Qy	1191	CATTTCAGATACCTCTCATACAGTGCACCCCTGCTTAAAGAAATGAAACCTATGCAATATGA	1251
Db	1191	CATTTCAGATACCTCTCATACAGTGCACCCCTGCTTAAAGAAATGAAACCTATGCAATATGA	1251
Qy	1251	CATCCACAGCGCTCGGTTAAATTAAAGGGGTGATCACCAGTTTCATATAATTTTCCCTTAT	1311
Db	1251	CATCCACAGCGCTCGGTTAAATTAAAGGGGTGATCACCAGTTTCATATAATTTTCCCTTAT	1311
Qy	1311	AAAAGGATTTGTGGCCACAGTGCATAGTGTTCATGCTGTAAATCCACAGATTTGGG	1366
Db	1311	AAAAGGATTTGTGGCCACAGTGCATAGTGTTCATGCTGTAAATCCACAGATTTGGG	1366
RESULT 14			
LOCUS	AX230123	1104 bp	DNA linear PAT 11-SEP-2001
DEFINITION	Sequence 10 from Patent WO0162797.		
ACCESSION	AX230123		
VERSION	AX230123.1	GI:15592145	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Vogel, G., Wood, L.S., Parodi, L.A. and Lind, P.		
TITLE	Novel g protein-coupled receptors		
JOURNAL	Patent: WO 0162797-A 10 30-AUG-2001;		
PHARMACIA	PHARMACIA & UPJOHN COMPANY (US)		
FEATURES			
source	1..1104		
	location/Qualifiers		
	1..1104		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	18.7%;	Score 326;	DB 6; Length 1104;
Best Local Similarity	100.0%;	Pred. No. 1.7e-177;	
Matches 326;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	248	CGACGCGCGGCGGAGCTGCGCTGCTGACTCAACTTCTGCGCGGACCTGCTCTTC	307
Db	27	CGACGCGCGGCGGAGCTGCGCTGCTGACTCAACTTCTGCGCGGACCTGCTCTTC	86
Qy	308	ATACAGGCTATCCTCTGGTGTGGCGCGGTGCGCTGAGCTGAGGCTGTGCTGTGGGCCCC	367

	Db	87	AATGACGGTATTCCTCTGTGTCTTGCCCGTGGCTGGACTGAAGGCTGGCTGTGGGCC	146
	Qy	368	GTTGGCTGCACCTGTCTTTCTTACGTATGACCCTGAGCGGACGTCATCCTTCA	427
	Db	147	GTTGGCTGCACCTGTCTTTCTTACGTATGACCCTGAGCGGACGTCATCCTTCA	206
	Qy	428	CTGGCCGGGCTGACGCTTGAGAGCATGTGTGTGATCTGTGACCTCTGACAGCCGGGGTGTGG	487
	Db	207	CTGGCCGGGCTGACGCTTGAGAGCATGTGTGTGATCTGTGACCTCTGACAGCCGGGGTGTGG	266
	Qy	488	GATCCTGTGGCGCGGCGGCGGCGAGTGTGTGTGAGCGCTCATCTGGAGCATATTGGCGGCTC	547
	Db	267	GATCCTGTGGCGCGGCGGCGGCGAGTGTGTGTGAGCGCTCATCTGGAGCATATTGGCGGCTC	326
	Qy	548	GCCGCTGTGCTCTCTGTGTGTCTTCTT	573
	Db	327	GCCGCTGTGCTCTCTGTGTGTCTTCTT	352
	RESULT 15			
	BD212934			
	LOCUS	BD212934	300 bp	DNA linear
	DEFINITION	Novel human genes and gene expression products II.	PAT 17-JUL-2003	
	ACCESSION	BD212934		
	VERSION	BD212934.1	GI:33022704	
	KEYWORDS	JP 2002519000-A/1076.		
	SOURCE	Homo sapiens (human)		
	ORGANISM	Homo sapiens		
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	AUTHORS	Mammalia; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.		
		1 (bases) to 300)		
		Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klingner,J.S.,		
		Reinhard,C., Giese,K., Randazzo,F., Kennedy,G.C., Pot,D.,		
		Kassam,A., Lamson,G., Drmanac,R., Ckvenjnikov,R., Dickson,M.,		
		Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V., Jones,L.W.		
		and Craih,B.S.		
	TITLE	Novel human genes and gene expression products II		
	JOURNAL	Patent: JP 2002519000-A 1076 02-JUL-2002;		
	COMMENT	CHIRON CORP. HYSEQ INC		
		OS Homo sapiens (human)		
		PN JP 2002519000-A/1076		
		PD 02-JUL-2002		
		PF 28-JAN-1999 JP 2000556580		
		PR 28-JAN-1998 US 60/072910,24-FEB-1998 US 60/075954 PR		
		31-MAR-1998 US 60/080114,03-APR-1998 US 60/080515 PR		
		03-APR-1998 US 60/080666,21-OCT-1998 US 60/105234 PR		
		28-OCT-1998 US 60/105877		
		PI LOUIS T WILLIAMS,JAIME ESCOBEDO,MICHAEL A INNIS,PABLO PI		
		DOMINEZ GARCIA,		
		PI JULIE SUDDUTH KLINGER,CHRISTOPH REINHARD,KLAUSE GIESE,FILIPPO		
		PI RANAZZO,		
		PI GIULIA C KENNEDY,DAVID POT,ALTF KASSAM,GEORGE LAMSON,RADOUJE		
		PI DRMANAC,		
		PI RADOMIR CRKVENJAKOV,MARK DICKSON,SNEZANA DRMANAC,IVAN LABAT,		
		PI DENA LESHKOWITZ,DAVID KITTA,VERONICA GARCIA,LEE WILLIAM JONES,		
		PI BIRGIT STACHE CRAIN		
		PC C12N15/09,C12N15/09,C07K14/47,C07K14/82,C07K16/18,C12N1/15, PC		
		C12N1/19'		
		PC C12N1/21,C12N5/10,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC		
		Novel human genes and gene expression products II FH key		
		Location/Qualifiers		
	FT	source	1..300	
		/organism='Homo sapiens (human)'.		
	FEATURES			
	Source		1..300	
		/organism='Homo sapiens (human)'		
		/mol_type='genomic DNA'		
		/db_xref='taxon:9606'		
	ORIGIN			
	Query Match	17.2%; Score 300; DB 6; Length 300;		
	Best Local Similarity	100.0%; Pred No. 2,1e-162;		

```
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 TGCTAATTCAGCCCTAAAGCCCATCCCTACACATGACACTGTGACGAATGAGTGA 1035
Db 1 TGCTAATTCAGCCCTAAAGCCCATCCCTACACATGACACTGTGACGAATGAGTGA 60
QY 1036 GAAAAATTTTGTGCTTCTGTGTTCCGAGAAAAGGAGCCATTTTAAACAGACATCTGT 1095
Db 61 GAAAAATTTTGTGCTTCTGTGTTCCGAGAAAAGGAGCCATTTTAAACAGACATCTGT 120
QY 1096 CAAAAAGAAATGACTTGTGATTTATTTTGGCTAATTTTCTTTATAGCCGAGTTTCTCAC 1155
Db 121 CAAAAAGAAATGACTTGTGATTTATTTTCTGCTAATTTTCTTTATAGCCGAGTTTCTCAC 180
QY 1156 ACCTGGGAGAGCTGTGGATGCTTTTAAACAGAGTTTCATTTCCAGTACCCTCCATCAGTGC 1215
Db 181 ACCTGGGAGAGCTGTGGATGCTTTTAAACAGAGTTTCATTTCCAGTACCCTCCATCAGTGC 240
QY 1216 ACCCTGCTTTAAGAAAAATGAACTATGCAAAATAGACATCCACAGCGTCGGTAAATTAAAG 1275
Db 241 ACCCTGCTTTAAGAAAAATGAACTATGCAAAATAGACATCCACAGCGTCGGTAAATTAAAG 300
```

Search completed: October 1, 2004, 01:41:55
Job time : 6621.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 19:01:32 ; Search time 635.881 Seconds
(without alignments)
11644.646 Million cell updates/sec

Title: US-10-077-698-2

Perfect score: 1743
Sequence: 1 tccgcagctagctctagacgcg.....aaaaaaaaaagggcg 1743

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneseqn19808:*\n2: geneseqn19908:*\n3: geneseqn20008:*\n4: geneseqn2001a8:*\n5: geneseqn2001b8:*\n6: geneseqn2002a8:*\n7: geneseqn2003a8:*\n8: geneseqn2003b8:*\n9: geneseqn2003c8:*\n10: geneseqn2004a8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1743	100.0	1743	3	AAAG4346
2	1743	100.0	1743	6	ABQ81226 Human 142
3	1328	76.2	1458	4	AAI66039 Human G P
4	1109	63.6	1160	7	ABX10627 DNA encod
5	1086	62.3	1086	6	ABST73343 CDNA enc
6	1081	62.0	1321	4	ABD08854 Human G-P
7	1035	59.4	1086	6	ABST73398 DNA encod
8	1035	59.4	1086	7	ACA60998 CDNA enc
9	1035	59.4	1086	8	ADBA47641 Human CDN
10	984	56.5	1086	6	ABST73399 DNA encod
11	926	53.1	1086	6	ABT04869 Human G P
12	913	52.4	1742	3	AAZ49745 Human 142
13	746	42.8	746	5	ABA21432 Human ner
14	586	33.6	632	7	ACA61000 Novel hum
15	586	33.6	632	8	ADBA47644 Human CDN
16	559	32.1	1066	9	ADCB7012 Human GPC
17	550	31.6	3173	4	AAZ82670 Human 7TM
18	425	24.4	933	9	ADCI2755 Human GPC
19	391	22.4	1737	7	ABZ42599 Human G P
20	337	19.3	388	4	ABA08489 Human GPC
21	326	18.7	1104	4	AAZ42815 Human G P
22	300	17.2	300	2	AAZ13607 Human gen
23	298	17.1	760	2	AAZ99045 Human val

24	198	11.4	300	2	AAZ98319 Human can
25	143	8.2	300	2	AAZ98320 Human can
26	131	7.5	638	6	AAZ62901 Human G P
27	111	6.4	426	4	AAZ42838 Human G P
28	67	3.8	3955	4	AAK66108 Human imm
29	67	3.8	11684	4	AAK66109 Human imm
30	67	3.8	12458	4	AAK66107 Human imm
31	66	3.8	700	4	AAH92222 Human inf
32	66	3.8	25871	3	AAH92222 Human inf
33	65	3.7	178896	6	ABQ88146 Human ost
34	58	3.3	162450	3	AAZ86967 Retinobla
35	57	3.3	80	7	ACA61001 Novel hum
36	57	3.3	80	8	ADBA47645 Human GPC
37	57	3.3	19553	8	ABK42538 Genomic s
38	57	3.3	19553	8	ABK42538 Genomic s
39	55	3.2	319	7	ABZ71836 Connectiv
40	55	3.2	598	4	AAK66647 Human can
41	55	3.2	14970	4	AAZ42037 Genomic s
42	55	3.2	14976	4	AAZ42038 Genomic s
43	55	3.2	27827	9	ADA02810 Human Fib
44	55	3.2	27827	9	ADB72548 Human CA
45	55	3.2	27828	9	ADC85290 Human gen

ALIGNMENTS

RESULT 1	AAA64346	standard; DNA; 1743 BP.
XX	AAA64346	
AC	AAA64346	
XX	20-DEC-2000	(first entry)
DT	20-DEC-2000	(first entry)
XX	DNA encoding a human G-protein coupled receptor designated 14273.	
DE	DNA encoding a human G-protein coupled receptor designated 14273.	
XX	Human, G-protein coupled receptor; receptor-mediated disorder;	
KM	14273 receptor; cardiovascular disease; congestive heart failure;	
KW	cardiac myocyte hypertrophy; ss.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	location/Qualifiers
FT	CDS	44..1129
FT	FT	/*tag= a
XX	FT	/product= "G-protein coupled receptor 14273"
XX	NO200050596-A2.	
XX	31-AUG-2000.	
PD	31-AUG-2000.	
XX	28-FEB-2000; 2000MO-US005068.	
PF	28-FEB-2000; 2000MO-US005068.	
XX	26-FEB-1999; 99US-00261599.	
PR	08-DEC-1999; 99US-00456455.	
XX	(MILL-) MILLENIUM PHARM INC.	
PA	(MILL-) MILLENIUM PHARM INC.	
XX	Gluckemann MA, Tsai F;	
PI	Gluckemann MA, Tsai F;	
XX	WPI, 2000-587184/55.	
DR	WPI, 2000-587184/55.	
DR	P-PSDB; AAB08538.	
XX	14273 receptor polynucleotides and polypeptides, useful in the diagnosis	
PT	and treatment of receptor-mediated disorders such as cardiovascular	
PT	diseases.	
XX	Claim 3, Fig 1, 105pp; English.	
PS	The present sequence encodes a human G-protein coupled receptor,	
XX	designated 14273 receptor. The G-protein coupled receptor 14273	
CC	polypeptide is used to produce antibodies, in drug screening assays, and	

CC to screen for agonist and antagonists. The antibodies are used for
CC diagnostic applications, and to assess abnormal tissue distributions or
CC abnormal expression during development. The antibodies are also useful
CC for tissue typing, and in forensic identification, as well as for
CC inhibiting receptor function. The 14273 polynucleotides can be used to
CC express the protein, to detect 14273 mRNA, to detect genetic alterations
CC in the 14273 gene, and as a source of primers and probes. The 14273
CC polynucleotides and polypeptides are also useful as a target for
CC diagnosis and treatment of receptor-mediated disorders, especially
CC cardiovascular diseases such as congestive heart failure caused by
CC cardiac myocyte hypertrophy
XX

Sequence 1743 BP; 380 A; 479 C; 451 G; 433 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 3; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TCCGACTAGTCTAGACCCGTCGCGGCGCCGACGCGCCGGGAATGTCCTGATGCGC 60
DB 1 TCCGACTAGTCTAGACCCGTCGCGGCGCCGACGCGCCGGGAATGTCCTGATGCGC 60
QY 61 GCGGCGACGCGGCGACCGCCCTTTCGCAAGCTTGAGCAAGCAACCGACCCGCTTTC 120
DB 61 GCGGCGACGCGGCGACCGCCCTTTCGCAAGCTTGAGCAAGCAACCGACCCGCTTTC 120
QY 121 CTTCTTCCGAGCTCAAGGCGACGACCGGCTGTGCGCGGCGGAGCAACCGT 180
DB 121 CTTCTTCCGAGCTCAAGGCGACGACCGGCTGTGCGCGGCGGAGCAACCGT 180
QY 181 GCTGAGTCTATCTTTGCAAGTGTGCTGCGGCAAGTGTGCGCCCTGTGCTGTGCGC 240
DB 181 GCTGAGTCTATCTTTGCAAGTGTGCTGCGGCAAGTGTGCGCCCTGTGCTGTGCGC 240
QY 241 GCGGCGACGACCGCGCGCGCGGAGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 GCGGCGACGACCGCGCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 GCTCTTATAGAGCTATCCCTCTGTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCT 360
DB 301 GCTCTTATAGAGCTATCCCTCTGTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GGGGCGCGTGTGCGGACGCTGCTCTTCTAGTATGACCTGAGCGGACGCTACCAT 420
DB 361 GGGGCGCGTGTGCGGACGCTGCTCTTCTAGTATGACCTGAGCGGACGCTACCAT 420
QY 421 CTTCAAGCTGTGCGCGGCTGAGCTGAGCGGATGTGTGCTGCTGCTGCTGCTGCTG 480
DB 421 CTTCAAGCTGTGCGCGGCTGAGCTGAGCGGATGTGTGCTGCTGCTGCTGCTGCTG 480
QY 481 CGTGCAGGCTCTGAGGCGGCGGCGGCGGAGTGTGCTGCGGCTGCTGCTGCTGCTGCT 540
DB 481 CGTGCAGGCTCTGAGGCGGCGGCGGCGGAGTGTGCTGCGGCTGCTGCTGCTGCTG 540
QY 541 GCGGCGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 GCGGCGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 CGCGGACGAGGAATTTGATGACATGATTTGCGCCCACTTCTGAGAGATCTC 660
DB 601 CGCGGACGAGGAATTTGATGACATGATTTGCGCCCACTTCTGAGAGATCTC 660
QY 661 GTGGGATGTCTCTTTGTTACTTTGAAGTTCTGTGTGCGAGGATGTGATGATGAG 720
DB 661 GTGGGATGTCTCTTTGTTACTTTGAAGTTCTGTGTGCGAGGATGTGATGATGAG 720
QY 721 TTACTCCAAATTTTACAGATCAAAAGGATCAAGAGAGGCTCAAGTAAAGCTTGC 780
DB 721 TTACTCCAAATTTTACAGATCAAAAGGATCAAGAGAGGCTCAAGTAAAGCTTGC 780
QY 781 CTACTCGAAGGACCAAGATCGCGGTGTCCAGACGAGACTTCCGCTCTTCCGACCT 840
DB 781 CTACTCGAAGGACCAAGATCGCGGTGTCCAGACGAGACTTCCGCTCTTCCGACCT 840
```

```
QY 841 CTTCTCTCATGAGTCTCTTCTTTCATCATGTGAGGCGCCCATCATCATCATCTCTCT 900
DB 841 CTTCTCTCATGAGTCTCTTCTTTCATCATGTGAGGCGCCCATCATCATCATCTCTCT 900
QY 901 CATCTGTATCAGAACTTCAAGCAAGACCTGTATCTGCGCGCTCTTCTTCTGAGT 960
DB 901 CATCTGTATCAGAACTTCAAGCAAGACCTGTATCTGCGCGCTCTTCTTCTGAGT 960
QY 961 GGTGCGCTTCACTTTGCTTAATTCAGCCCTTAACCCCATCTCTACAAACATGACGTG 1020
DB 961 GGTGCGCTTCACTTTGCTTAATTCAGCCCTTAACCCCATCTCTACAAACATGACGTG 1020
QY 1021 CAGGAATGATGAGAAATTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 CAGGAATGATGAGAAATTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 AACGACACATCTGTCAAAAGAAATGAATTTGATTAATTTTCTGCTTAATTTTCTTAT 1140
DB 1081 AACGACACATCTGTCAAAAGAAATGAATTTGATTAATTTTCTGCTTAATTTTCTTAT 1140
QY 1141 AGCGAATTTCTCACACCTGCGGAGCTGTGCGATCTTTTAAACAGTTCAATTCAGT 1200
DB 1141 AGCGAATTTCTCACACCTGCGGAGCTGTGCGATCTTTTAAACAGTTCAATTCAGT 1200
QY 1201 ACCCTCATGATGACACCTGCTTTTAAAGAAATGAACCTTATGCAATATGACATCAGC 1260
DB 1201 ACCCTCATGATGACACCTGCTTTTAAAGAAATGAACCTTATGCAATATGACATCAGC 1260
QY 1261 GTGCGTAATTAAGGCGGTATCAACCAAGTTTCAATATATTTTCCCTTATTAAGATTT 1320
DB 1261 GTGCGTAATTAAGGCGGTATCAACCAAGTTTCAATATATTTTCCCTTATTAAGATTT 1320
QY 1321 GTTGGCCAGGTGAGAGGCTTCAAGCTGTATCCAGAGTTTGGAGGCGTGAAGTGGT 1380
DB 1321 GTTGGCCAGGTGAGAGGCTTCAAGCTGTATCCAGAGTTTGGAGGCGTGAAGTGGT 1380
QY 1381 GGAATCACTGAGGTCAAGAGTGTGAGACCAACCTGACCAACATGTGTGAGACCCCGTCTC 1440
DB 1381 GGAATCACTGAGGTCAAGAGTGTGAGACCAACCTGACCAACATGTGTGAGACCCCGTCTC 1440
QY 1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
DB 1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
QY 1501 CTTGGAGGCTGAACCGAGGAATCTTGAACCTGGAAGGCGAGAGTGTGAGTGAACCG 1560
DB 1501 CTTGGAGGCTGAACCGAGGAATCTTGAACCTGGAAGGCGAGAGTGTGAGTGAACCG 1560
QY 1561 AGATCGGCAATGCACTCCAAACGAGGCAACAGAGTGAATCTCTTAATTAATTAATTA 1620
DB 1561 AGATCGGCAATGCACTCCAAACGAGGCAACAGAGTGAATCTCTTAATTAATTAATTA 1620
QY 1621 AAAAAAAGATTTGTTATGAGGCTCTTAAATGTGAACCTTTTATGTTGTTAATA 1680
DB 1621 AAAAAAAGATTTGTTATGAGGCTCTTAAATGTGAACCTTTTATGTTGTTAATA 1680
QY 1681 TGAATCAATTTAATTAATTAATTTAATTAATGATGCTGTCGCAAAAAAAGAGAGG 1740
DB 1681 TGAATCAATTTAATTAATTAATTTAATTAATGATGCTGTCGCAAAAAAAGAGAGG 1740
QY 1741 CGG 1743
DB 1741 CGG 1743
```

RESULT 2
AB081226
ID AB081226 standard; cDNA; 1743 BP.
XX AB081226;
AC
XX
DT 05-DEC-2002 (first entry)

XX Human 14273 nucleic acid, associated with metabolic disorder.
 DE Human; 14273; metabolic disorder; obesity; diabetes; anorexia; cachexia;
 KW anorectic; antidiabetic; anabolic; transgenic animal; gene therapy; gene;
 KM 88.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 44..1129
 FT CDS /*tag= a
 FT /product= "14273"
 XX
 XX MO200267868-A2.
 XX
 XX 06-SEP-2002.
 XX
 XX 26-FEB-2002; 2002MO-US006131.
 XX
 XX 26-FEB-2001; 2001US-0271655P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Gimeno R, Tsai F;
 XX WPI; 2002-698629/75.
 DR P-PSDB; ABB79906.
 XX
 XX
 XX Identifying a nucleic acid associated with a metabolic disorder, useful
 PT for diagnosing metabolic disorders, e.g. obesity, comprises contacting
 PT the sample with a probe comprising at least 25 contiguous nucleotides of
 PT the 14273 gene.
 XX
 XX Claim 1; Fig 1A-B; 95pp; English.
 PS
 XX The present sequence is that of a human nucleic acid, designated 14273,
 CC associated with metabolic disorders. 14273 molecules are expressed at
 CC high levels in adipose tissue, e.g. white adipose tissue and brown
 CC adipose tissue, as well as in pancreatic islets. They are upregulated
 CC during exposure to cold (i.e. under conditions that affect brown or white
 CC adipocyte metabolism) and downregulated in genetic models of obesity.
 CC 14273 knock-out mice, when fed a high-fat diet, gain more weight and have
 CC larger epididymal fat pads than wild-type mice. They also show increased
 CC levels of glucose and insulin upon fasting. A 14273 agonist may be
 CC beneficial to the treatment of obesity and/or type II diabetes by
 CC preventing fat accumulation on a high-fat diet and/or the increases in
 CC endogenous glucose production which occur in type II diabetes. The
 CC present invention provides 14273 nucleic acids, polypeptides and
 CC antibodies useful for the diagnosis and treatment of metabolic disorders
 CC including obesity, anorexia, cachexia and diabetes. Also provided are
 CC methods for identifying a subject having a metabolic disorder, for
 CC identifying a compound capable of modulating metabolic activity, methods
 CC for modulating metabolic activity or adipocyte activity (hyperplastic
 CC growth, hypertrophic growth or lipogenesis), methods for modulating
 CC lipogenesis or lipolysis in a subject, and a method for regulating
 CC endogenous glucose levels
 CC
 XX Sequence 1743 BP; 380 A; 479 C; 451 G; 433 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 1743; DB 6; Length 1743;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 CTTCTTCCGAGCTCAAGGCGACACACCGCTGCTGCTGCGCGGATGAGACACCT 180
 Qy 181 GCTGCTCATCTTTTCAAGTGTGCTGCTGCGCAAGTGTGCGCTGTGCTGTGCGC 240
 Db 181 GCTGCTCATCTTTTCAAGTGTGCTGCTGCGCAAGTGTGCGCTGTGCTGTGCGC 240
 Qy 241 GCGCGGACGACGCGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Db 241 GCGCGGACGACGCGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Qy 301 GCTCTTATCAGGCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Db 301 GCTCTTATCAGGCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Qy 361 GGGCGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 GGGCGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Qy 421 CTTCAAGCTGCGCGGCTGAGCTGAGCGGCTGAGCTGAGCTGAGCTGAGCTGAG 480
 Db 421 CTTCAAGCTGCGCGGCTGAGCTGAGCGGCTGAGCTGAGCTGAGCTGAGCTGAG 480
 Qy 481 GGTGCGGGGTCCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 Db 481 GGTGCGGGGTCCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 Qy 541 GGGCGGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Db 541 GGGCGGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Qy 601 GCGCGACCAAGAAATTTGATTTGACATGATTTGGCCCACTTCTTCTGAGAGAT 660
 Db 601 GCGCGACCAAGAAATTTGATTTGACATGATTTGGCCCACTTCTTCTGAGAGAT 660
 Qy 661 GTGGAGATGCTCTTTTGTATCTTGAACCTTGTGCTGCAAGATGCTGATTTGAT 720
 Db 661 GTGGAGATGCTCTTTTGTATCTTGAACCTTGTGCTGCAAGATGCTGATTTGAT 720
 Qy 721 TTATCCCAAAATTTTACAGATCAACAAAGCATCAAGAGAGCTCAAGCTGAGC 780
 Db 721 TTATCCCAAAATTTTACAGATCAACAAAGCATCAAGAGAGCTCAAGCTGAGC 780
 Qy 781 CTACTGCGAAGACCAAGATCCGCGGTGCTCCAGAGAGATTTCCGCTCTTCCG 840
 Db 781 CTACTGCGAAGACCAAGATCCGCGGTGCTCCAGAGAGATTTCCGCTCTTCCG 840
 Qy 841 CTTCTCTCATGAGTCTCTTCTTCAATCATGTGAGGCCCATCATCATCATCTCT 900
 Db 841 CTTCTCTCATGAGTCTCTTCTTCAATCATGTGAGGCCCATCATCATCATCTCT 900
 Qy 901 CATCTGATCAGAACTTCAAGCAAGACCTGATCTGAGCTGCTCTTCTTCTGAG 960
 Db 901 CATCTGATCAGAACTTCAAGCAAGACCTGATCTGAGCTGCTCTTCTTCTGAG 960
 Qy 961 GGTGCGCTTCAATTTGCTAATTCAGCCCTAAACCCCATCTCTACAACTGACAC 1020
 Db 961 GGTGCGCTTCAATTTGCTAATTCAGCCCTAAACCCCATCTCTACAACTGACAC 1020
 Qy 1021 CAGGAATGATGGAAGAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1021 CAGGAATGATGGAAGAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Qy 1081 AAGAGACATCTGTCAAAAGAAATGACTGTGATTAATTTGCTGCTGCTGCTTAT 1140
 Db 1081 AAGAGACATCTGTCAAAAGAAATGACTGTGATTAATTTGCTGCTGCTGCTTAT 1140
 Qy 1141 AGCGAGTTTCTCAACCTGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 1141 AGCGAGTTTCTCAACCTGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Qy 1201 ACCCTCATGATGACCTGCTTTAAGAAATGAACCTTATGCAATATGACATCAGC 1260

QY 557 CCTCTGCGCTCTTTCAGTGTCCGCAAGGCTCCCGGCGCGCAAGAAATT 616
 Db 541 CCTCTGCGCTCTTTCAGTGTCCGCAAGGCTCCCGGCGCGCAAGAAATT 600
 QY 617 TCGATTTGCACTGATTTGGCCCACTTCTGAGAGATCTTGAGATCTCTTTT 676
 Db 601 TCGATTTGCACTGATTTGGCCCACTTCTGAGAGATCTTGAGATCTCTTTT 660
 QY 677 GTTACTTTGAATCTTGTGTGCGCAGGACTGTGATTTGATGATCTTCAAAATT 736
 Db 661 GTTACTTTGAATCTTGTGTGCGCAGGACTGTGATTTGATGATCTTCAAAATT 720
 QY 737 CAGATCAAAAGGATCAAGAGAGGCTCAAGGTAAGCTTCTGAGAGGCTTCTGAGAGGCT 796
 Db 721 CAGATCAAAAGGATCAAGAGAGGCTCAAGGTAAGCTTCTGAGAGGCTTCTGAGAGGCT 780
 QY 797 CAGATCCGCTGTGCGCAGGACTTCCGCTCTTCCGCACTTCTCTCTCATGCTC 856
 Db 781 CAGATCCGCTGTGCGCAGGACTTCCGCTCTTCCGCACTTCTCTCTCATGCTC 840
 QY 857 TCCCTTTCATGATGAGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAGAGGCT 916
 Db 841 TCCCTTTCATGATGAGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAGAGGCT 900
 QY 917 TTCAGCAAGACCTGATCTGAGGCTTCTCTTCTGAGAGGCTTCTGAGAGGCT 976
 Db 901 TTCAGCAAGACCTGATCTGAGGCTTCTCTTCTGAGAGGCTTCTGAGAGGCT 960
 QY 977 GCTAATTCAGGCTTAAACCCCATCTTCAACATGACATGTCGAGATGAGTGAAG 1036
 Db 961 GCTAATTCAGGCTTAAACCCCATCTTCAACATGACATGTCGAGATGAGTGAAG 1020
 QY 1037 AAAATTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
 Db 1021 AAAATTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 QY 1097 AAAAGAAATGATCTGATTTTCTGCTAATTTTCTTATATGCGAGTTTCTGCA 1156
 Db 1081 AAAAGAAATGATCTGATTTTCTGCTAATTTTCTTATATGCGAGTTTCTGCA 1140
 QY 1157 CTTGGCAGCTGTGGCATGC 1176
 Db 1141 CTTGGCAGCTGTGGCATGC 1160

RESULT 5
 ABS7343
 ID ABS7343 standard; cDNA; 1086 BP.
 XX
 AC ABS7343;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE cDNA encoding human GPCR HF1948.
 XX
 KM Human; transmembrane receptor; G-protein coupled receptor; GPCR; allergy;
 KM hypertension; reflux disease; depression; migraine; schizophrenia; ulcer;
 KM psychiatric disorder; asthma; bronchospasm; anaesthesia;
 KM myocardial infarction; MI; stroke; glaucoma; anxiety;
 KM prostatic hyperplasia; epilepsy; prostate cancer; rhinitis; angina;
 KM prostatic hypertrophy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200268600-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 26-FEB-2002; 2002WO-US005625.
 XX
 PR 26-FEB-2001; 2001US-0271913P.
 XX

PA (AREN-) ARENA PHARM INC.
 XX
 PI L'aw CM, Chalmers DT, Behan DP, Maciejewski-Jenior D, Leonard JN;
 PI Lin I, Ortuno D;
 XX
 DR WPI, 2002-706980/76.
 DR P-PSDB; ABG95157.
 XX
 PT New human G-protein coupled receptor (GPCR), useful for screening agonist
 PT or inverse agonist compounds for treating diseases associated with GPCR.
 XX
 PS Claim 23; Page 122-123; 201pp; English.
 XX
 CC The present invention relates to transmembrane receptors, particularly
 CC endogenous human G-protein coupled receptors (GPCRs), mutant (non-
 CC endogenous) versions of the GPCRs, and the polynucleotide sequences
 CC encoding them. The GPCRs are useful for screening agonist or inverse
 CC agonist compounds for treating diseases associated with GPCR. Diseases
 CC that can be treated with such compounds include allergies, hypertension,
 CC reflux disease, depression, migraine, schizophrenia, ulcers, psychotic
 CC disorders, asthma, bronchospasm, anaesthesia, myocardial infarction (MI),
 CC stroke, glaucoma, prostatic hyperplasia, epilepsy, prostate cancer,
 CC anxiety, prostatic hypertrophy, rhinitis, and angina. The present
 CC sequence encodes an endogenous human GPCR
 XX
 SQ Sequence 1086 BP; 182 A; 348 C; 295 G; 261 T; 0 U; 0 Other;
 Query Match 62.3%; Score 1086; DB 6; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 44 ATGTCCCTGAATGCGCGGAGGAGCGGCGGAGCGGCTTGTGCGAGCTTGAAGAGCC 103
 Db 1 ATGTCCCTGAATGCGCGGAGGAGCGGCGGAGCGGCTTGTGCGAGCTTGAAGAGCC 60
 QY 104 AACCGACCGGCTTTTCCCTTCTTCCGAGTGAAGGCGGAGCCAGCGGCTGAGTGGCC 163
 Db 61 AACCGACCGGCTTTTCCCTTCTTCCGAGTGAAGGCGGAGCCAGCGGCTGAGTGGCC 120
 QY 164 GCGGTGAGACAACCGGCTGAGTCTATCTTGTGAGTGTGCTGAGGCAAGTGTGC 223
 Db 121 GCGGTGAGACAACCGGCTGAGTCTATCTTGTGAGTGTGCTGAGGCAAGTGTGC 180
 QY 224 GCCCTGTGTGTGTGCGCGGCGGAGCGCGGCGGAGCTGCTGCTGTAAC 283
 Db 181 GCCCTGTGTGTGTGCGCGGCGGAGCGCGGCGGAGCTGCTGCTGTAAC 240
 QY 284 CTCTTGTGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 343
 Db 241 CTCTTGTGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 300
 QY 344 ACTGAGGCTGTGCTGAGGCGGCTTGTGCTGAGCACTGCTTCTTCTTCTTCTTCTTCTTCT 403
 Db 301 ACTGAGGCTGTGCTGAGGCGGCTTGTGCTGAGCACTGCTTCTTCTTCTTCTTCTTCTTCT 360
 QY 404 AGCGGACGCTACCATCTTCAAGCTGCGCGGCTGAGCGGCTGAGCGCATGTGTGATC 463
 Db 361 AGCGGACGCTACCATCTTCAAGCTGCGCGGCTGAGCGGCTGAGCGCATGTGTGATC 420
 QY 464 GTGCACTTGAAGCGGCGGCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 523
 Db 421 GTGCACTTGAAGCGGCGGCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
 QY 524 CTATCTGGGCTATTTGCGGCTGCGGCTGCTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 583
 Db 481 CTATCTGGGCTATTTGCGGCTGCGGCTGCTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 540
 QY 584 CCGCAACGCTTCCCGGCGGCGGAGCAAGAAATTTGCAATTTGCACTGATTTGGCCCAAC 643
 Db 541 CCGCAACGCTTCCCGGCGGCGGAGCAAGAAATTTGCAATTTGCACTGATTTGGCCCAAC 600
 QY 644 ATTCTGGAAGATCTGTGAGATGTCTTTTGTATCTTGAACCTTGTGTGCGAGGA 703

```

Db      601 ATTCTGAGAGAGATCTGTGGAGATGCTCTTTTGTACTTTGAACCTCTTGTCAGAGA 660
Qy      704 CTGTGCTATTGTGATCACTTACTCTCAAAATTTTACAGATCAACAAAGCATCAGAGAGAG 763
Db      661 CTGTGCTATTGTGATCACTTACTCTCAAAATTTTACAGATCAACAAAGCATCAGAGAGAG 720
Qy      764 CTACAGGTAAAGCTGGCTCTCTGAGAGAGCAGACAGATCCGCGTGTCCGACGAGACTTC 823
Db      721 CTACAGGTAAAGCTGGCTCTCTGAGAGAGCAGACAGATCCGCGTGTCCGACGAGACTTC 780
Qy      824 CGGCTCTTCCGACCCCTCTTCTCTCATGAGTCTCTTCTTCAATCATGTGAGAGCCCATC 883
Db      781 CGGCTCTTCCGACCCCTCTTCTCTCATGAGTCTCTTCTTCAATCATGTGAGAGCCCATC 840
Qy      884 ATATATCAATCACTCTCTCTCATCTGATCCAGAACTTCAAGAGACCTGTGATCTGGCCG 943
Db      841 ATATATCAATCACTCTCTCTCATCTGATCCAGAACTTCAAGAGACCTGTGATCTGGCCG 900
Qy      944 TCCCTCTTCTTGGGGTGGTGGGCTTCAATTTGCTAATTCAGGCCCTAAACCCCATCTC 1003
Db      901 TCCCTCTTCTTGGGGTGGTGGGCTTCAATTTGCTAATTCAGGCCCTAAACCCCATCTC 960
Qy      1004 TACAACATGACACTGTGTCAGAGAAATGAGTGAAGAAATTTTGTGCTCTTCTGTTCCCA 1063
Db      961 TACAACATGACACTGTGTCAGAGAAATGAGTGAAGAAATTTTGTGCTCTTCTGTTCCCA 1020
Qy      1064 GAAAGGAGACCAATTTTACAGACACATCTGTCAAAAGAAATGACCTTGTGATTTTCT 1123
Db      1021 GAAAGGAGACCAATTTTACAGACACATCTGTCAAAAGAAATGACCTTGTGATTTTCT 1080
Qy      1124 GGCTAA 1129
Db      1081 GGCTAA 1086

RESULT 6
AAD08854
ID      AAD08854 standard, cDNA, 1321 BP.
XX
XX
AC      AAD08854;
XX
DT      04-SEP-2001 (first entry)
XX
DE      Human G-protein coupled receptor-20 (GCREC-20) cDNA.
XX
XX      Human, G-protein coupled receptor-20; GCREC-20; gene therapy; cirrhosis;
XX      transgenic animal; proliferative disorder; actinic keratosis;
XX      hepatitis nephrotoxic; cancer; breast; bladder; bone marrow; brain;
XX      uterus; leukaemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy;
XX      stroke; neurological disorder; Alzheimer's disease; Parkinson's disease;
XX      nausea; Huntington's disease; multiple sclerosis; dementia;
XX      angina pectoris; central nervous system disorder;
XX      cardiovascular disorder; hypertension; atherosclerosis;
XX      congestive heart failure; gastrointestinal disorder; dysphagia;
XX      peptic oesophagitis; spasm; gastritis; anorexia; pyrosis; pancreatitis;
XX      Crohn's disease; diarrhoea; autoimmune disorder; anaemia;
XX      inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;
XX      Addison's disease; allergy; asthma; diabetes mellitus; antihypertoid;
XX      atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;
XX      rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;
XX      metabolic disorder; obesity; nootropic; protozoacide; virucide; se.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      30..1115
FT      /tag= a
FT      /product= "Human GCREC-20 protein"
XX
XX      MO200142288-A2.
XX      14-JUN-2001.

```

```

PF      07-DEC-2000; 2000WO-US033382.
XX
XX      10-DEC-1999; 99US-0172852P.
PR      22-DEC-1999; 99US-017132P.
PR      14-JAN-2000; 2000US-0176148P.
PR      21-JAN-2000; 2000US-0177311P.
XX
XX      (INCYT-) INCYTE GENOMICS INC.
XX
XX      Burford N, Baughn MR, Au-Young J, Yang J, Lu DM, Reddy R.
DR      WPI; 2001-381635/40.
DR      P-PSDB; AAB04564.
XX
XX      New human G-protein coupled receptor polypeptides for diagnosing,
PT      preventing, and treating cell proliferative, neurological,
PT      cardiovascular, gastrointestinal, autoimmune and metabolic disorders.
XX
XX      Claim 5; Page 167; 175pp; English.
XX
XX      The present sequence is human G-protein coupled receptor-20 (GCREC-20)
CC      cDNA. GCREC is useful in somatic or germ-line gene therapy to correct a
CC      genetic deficiency, to express a conditionally lethal gene product and to
CC      express a protein which affords protection against intracellular
CC      parasites and also for diagnosis of disorders associated with expression
CC      of GCREC. GCREC is also useful for generating hybridisation probes useful
CC      in mapping the naturally occurring genomic sequences and to create
CC      knock-in humanised animals (pigs) or transgenic animals (mice or rats) to
CC      model human diseases. GCREC is used to diagnose, prevent and treat
CC      proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,
CC      hepatitis and cancer), cancer (breast, bladder, bone marrow, brain,
CC      uterine cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)
CC      neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's,
CC      Parkinson's disease, multiple sclerosis; dementia and other central
CC      nervous system disorders); cardiovascular disorders (angina pectoris,
CC      hypertension, atherosclerosis, congestive heart failure);
CC      gastrointestinal disorders (dysphagia, peptic oesophagitis, oesophageal
CC      spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina,
CC      pyrosis, pancreatitis, Crohn's disease, diarrhoea); autoimmune/
CC      inflammatory disorders (acquired immunodeficiency syndrome (AIDS),
CC      Addison's disease, allergies, anaemia, asthma, diabetes mellitus, atopic
CC      dermatitis, glomerulonephritis, Grave's disease, osteoarthritis,
CC      psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal,
CC      parasitic, protozoal and helminthic infections) and metabolic disorders
CC      (obesity, osteoporosis, viral infections)
XX
XX      Sequence 1321 BP; 243 A; 404 C; 348 G; 326 T; 0 U; 0 Other;
SQ
Query Match      62.0%; Score 1081; DB 4; Length 1321;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1181; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      32 CAGGCGCGGAGATGTCCTGTAATGCGCGGAGCGGAGCGGAGCCCTTGCGGAGC 91
Db      18 CAGGCGCGGAGATGTCCTGTAATGCGCGGAGCGGAGCGGAGCCCTTGCGGAGC 77
Qy      92 CTGAGACCAACCAACCGACCCGCTTCTCTTCTCGACGTCAAGGAGCAACACCG 151
Db      78 CTGAGACCAACCAACCGACCCGCTTCTCTTCTCGACGTCAAGGAGCAACACCG 137
Qy      152 CTGTGTCTGCGCGCGGTGAGACAAACGTCGTGTCTCATCTTTGACAGTGTGCTGCTG 211
Db      138 CTGTGTCTGCGCGCGGTGAGACAAACGTCGTGTCTCATCTTTGACAGTGTGCTGCTG 197
Qy      212 GGCAAGTGTGCGCGCTGTGTCGTGTGCGCGCGGAGAGCGCGGAGCGGAGCTGCTGCG 271
Db      198 GGCAAGTGTGCGCGCTGTGTCGTGTGCGCGCGGAGAGCGCGGAGCGGAGCTGCTGCG 257
Qy      272 CTGTACTCAACTCTTCTGTGCGGAGACTGCTTTCATGAGCGCTAATCCCTGTGTCGTG 331
Db      258 CTGTACTCAACTCTTCTGTGCGGAGACTGCTTTCATGAGCGCTAATCCCTGTGTCGTG 317
Qy      332 GCCGTGCGCTGACATGAGGCGTGTGCTGTGCGCGCGGCTTGCCCTGCAACTGCTCTTCTAC 391

```



```

QY 44 ATGTCCCTGAATACGCGCGGAGCGAGCGCCCTTGGCAGCCTGAGCAAGCC 103
DB 1 ATGTCCCTGAATACGCGCGGAGCGAGCGCCCTTGGCAGCCTGAGCAAGCC 60
QY 104 AACCGCACCCGCTTTTCTTTCTTCCGACGTCAAGGGGCAACAACCGGCTGGTGGCC 163
DB 61 AACCGCACCCGCTTTTCTTTCTTCCGACGTCAAGGGGCAACAACCGGCTGGTGGCC 120
QY 164 GCGGTGGAGCAACCGGCTGGTGGTCAATTTTGGCAAGTGGCTGGGCAACGTGTG 223
DB 121 GCGGTGGAGCAACCGGCTGGTGGTCAATTTTGGCAAGTGGCTGGGCAACGTGTG 180
QY 224 GCCCTGGTGGTGGGCGCGGACGACGCGCGCGCGGACGCTGGCTGCTGTAATCAAC 283
DB 181 GCCCTGGTGGTGGGCGCGGACGACGCGCGCGCGGACGCTGGCTGCTGTAATCAAC 240
QY 284 CTCTTCTGGCGGACCTGCTCTTCTATCAAGCTATCCCTTGGTGGCTGGCCGTGG 343
DB 241 CTCTTCTGGCGGACCTGCTCTTCTATCAAGCTATCCCTTGGTGGCTGGCCGTGG 300
QY 344 ACTGAGGCGCTGGGCTGGGCGCGGCTGGGCAACCTGCTCTTCTATCAAGTGAACCTG 403
DB 301 ACTGAGGCGCTGGGCTGGGCGCGGCTGGGCAACCTGCTCTTCTATCAAGTGAACCTG 360
QY 404 AGCGGACGCTCAACATCTCAGCTGGCGCGGCTGAGCTGGAGCGCATGCTGTGATC 463
DB 361 AGCGGACGCTCAACATCTCAGCTGGCGCGGCTGAGCTGGAGCGCATGCTGTGATC 420
QY 464 GTGCACTGCAAGCGCGGCTGGGAGTCTGGGCGCGGCGCGGAGTGTGCTGGCG 523
DB 421 GTGCACTGCAAGCGCGGCTGGGAGTCTGGGCGCGGCGCGGAGTGTGCTGGCG 480
QY 524 CTGATCTGGGGGCTATTCGGGGGTGGCGGCTGGCTCTGCGGCTCTTTCAGTGTG 583
DB 481 CTGATCTGGGGGCTATTCGGGGGTGGCGGCTGGCTCTGCGGCTCTTTCAGTGTG 540
QY 584 CCGCAACGCTCCCGCGCGCGGACGACGAAATTTGATGACATGATTTGGCCACC 643
DB 541 CCGCAACGCTCCCGCGCGCGGACGACGAAATTTGATGACATGATTTGGCCACC 600
QY 644 ATTCTGAGAGATCTGTGGGATGTCTTTTGTATCTTTGAACTTCTGGTGGCAAGA 703
DB 601 ATTCTGAGAGATCTGTGGGATGTCTTTTGTATCTTTGAACTTCTGGTGGCAAGA 660
QY 704 CTGTGATTTGATCACTTCTCCAAATTTTACAGATCAAGAGATCAAGAGAGAG 763
DB 661 CTGTGATTTGATCACTTCTCCAAATTTTACAGATCAAGAGATCAAGAGAGAG 720
QY 764 CTGACGCTAAGCTGGCTGCTGAGAGGACACAGATCGCGTGTCCAGCAGACTTC 823
DB 721 CTGACGCTAAGCTGGCTGCTGAGAGGACACAGATCGCGTGTCCAGCAGACTTC 780
QY 824 CGGCTTTCGCAACCTTCTCTCTCATGCTCTTCTTTCATCATGTGAGGCCATC 883
DB 781 CGGCTTTCGCAACCTTCTCTCTCATGCTCTTCTTTCATCATGTGAGGCCATC 840
QY 884 ATGATCAACATCTCCCTCATCTGATCCAGAACTTCAAGAGAGCTGGTATGTGGCG 943
DB 841 ATGATCAACATCTCCCTCATCTGATCCAGAACTTCAAGAGAGCTGGTATGTGGCG 900
QY 944 TCCTCTTCTTGGGTGGTGGCTTCAATTTGCTAATTCAGCCCTTAACCCCATCTC 1003
DB 901 TCCTCTTCTTGGGTGGTGGCTTCAATTTGCTAATTCAGCCCTTAACCCCATCTC 960
QY 1004 TACAACATGACATCTGTGAGAGATGATGAGAAATTTTGGTGTGTTCCCA 1063
DB 961 TACAACATGACATCTGTGAGAGATGATGAGAAATTTTGGTGTGTTCCCA 1020
QY 1064 GAAAGGAGAGCATTTTAAAGAGACATCTGTCAAAAGAAATGACTGTGATTAATTTCT 1123
DB 1021 GAAAGGAGAGCATTTTAAAGAGACATCTGTCAAAAGAAATGACTGTGATTAATTTCT 1080
QY 1124 GGCTAA 1129

```

```

DB 1081 GGCTAA 1086
|||||
RESULT 11
ABT04869
ID ABT04869 standard; cDNA; 1086 BP.
XX
AC ABT04869;
XX
DE 11-OCT-2002 (first entry)
XX
DE Human G protein coupled receptor hRUP31 coding sequence.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001MO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0255366P.
PR 20-FEB-2001; 2001US-0270266P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282356P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282359P.
PR 14-MAY-2001; 2001US-0290917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
XX (AREN-) ARENA PHARM INC.
XX
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
XX WPI; 2002-566565/60.
XX
DR P-PSDB; ABJ04071.
XX
PT Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
XX Claim 15; Page 61; 84pp; English.
XX
XX The present invention provides the protein and coding sequences of
XX several human G-protein coupled receptors (GPCRs). These can be used in
XX the identification of candidate compounds as receptor agonists or inverse
XX agonists having applicability as therapeutic agents. The present sequence
XX is a GPCR coding sequence of the invention
XX
SQ Sequence 1086 BP; 185 A; 346 C; 295 G; 260 T; 0 U; 0 Other;
Query Match 53.1%; Score 926; DB 6; Length 1086;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 51 CTGAATCGCGGAGCGAGCGGCGGAGCGGCTTGGCAGCCTGAGCAAGCAACCGCA 110
DB 8 CTGAATCGCGGAGCGGAGCGGCGGAGCGGCTTGGCAGCCTGAGCAAGCAACCGCA 67
QY 111 CCGGCTTTCCTTCTCTCGGACGTCAAGGCGCAACCGGCTGGTGTGCGCGGCTGG 170
DB 68 CCGGCTTTCCTTCTCTCGGACGTCAAGGCGCAACCGGCTGGTGTGCGCGGCTGG 127
QY 171 AGACAACGCTGCTGCTATCTTTCAGATGTGCTGCTGGGCAAGTGTGCGCCCTGG 230
DB 128 AGACAACGCTGCTGCTATCTTTCAGATGTGCTGCTGGGCAAGTGTGCGCCCTGG 187

```

OY		221	TGCTGGTGGCGGCGCAGACGACCCGGGGGCATCGCCCTGTACTCAACCTTTCT	290
Db		188	TGCTGGTGGCGGCGCAGACGACCCGGGGGCATCGCCCTGTACTCAACCTTTCT	247
OY		291	GCGCGGACCTGTCTTTCATCAGCGCATCCCTGTGGTGCGGCCGTGGACTGAAG	350
Db		248	GCGCGGACCTGTCTTTCATCAGCGCATCCCTGTGGTGCGGCCGTGGACTGAAG	307
OY		351	CCTGGCTGCTGGGCCCCGTTGCTGCGACCTGCTCTTCTACGTGATGACCCTGACCGCA	410
Db		308	CCTGGCTGCTGGGCCCCGTTGCTGCGACCTGCTCTTCTACGTGATGACCCTGACCGCA	367
OY		411	GCGTACCATCCTCACAGCTGGGCGGGGTACAGCTGTGAGGCGGATGGTGTGACCTGTCAC	470
Db		368	GCGTACCATCCTCACAGCTGGGCGGGGTACAGCTGTGAGGCGGATGGTGTGACCTGTCAC	427
OY		471	TGCACGCGGCGGTGGGGGTCTTGGGCGGCGGGCGGCGAGTGTGCTGGCGCTCATCT	530
Db		428	TGCACGCGGCGGTGGGGGTCTTGGGCGGCGGGCGGCGAGTGTGCTGGCGCTCATCT	487
OY		531	GGGGCTATTCCGGCGGTGCGCGGCTTGCCTCTGCGGTCTTCTTTCAGATCGTCCGCAAC	590
Db		488	GGGGCTATTCCGGCGGTGCGCGGCTTGCCTCTGCGGTCTTCTTTCAGATCGTCCGCAAC	547
OY		551	GGCTCCCCGGCGCCCAACGAGAAATTTGCACTGTATTTGGGCCACCATTCCTCG	607
Db		548	GGCTCCCCGGCGCCCAACGAGAAATTTGCACTGTATTTGGGCCACCATTCCTCG	607
OY		651	GAGAGATCGTGGGAGTGTCTTTTGTGTAATTAATCTTGGTGGCCAGGACTGTGTCA	710
Db		608	GAGAGATCTCGTGGAGTGTCTTTTGTGTAATTAATCTTGGTGGCCAGGACTGTGTCA	667
OY		711	TTGTGATCAGTTACTCCAATAATTTTACAGATCACAAAGGCATCAAGAAGAGCTCACGG	770
Db		668	TTGTGATCAGTTACTCCAATAATTTTACAGATCACAAAGGCATCAAGAAGAGCTCACGG	727
OY		771	TAAGCTTGCGCTAATCTGCGAGAGCCACAGATCCGGGTGCCAGACGACTTCGGGCTCT	830
Db		728	TAAGCTTGCGCTAATCTGCGAGAGCCACAGATCCGGGTGCCAGACGACTTCGGGCTCT	787
OY		831	TCCGACCCCTCTTCTCCTCATAGTGTCTCTTTCATCATGTGAGGCCCATCATCATCA	890
Db		788	TCCGACCCCTCTTCTCCTCATAGTGTCTCTTTCATCATGTGAGGCCCATCATCATCA	847
OY		891	CCATCTCTCTCATCTGTATCCAGAACCTTCAAGCAGACCTGTGATCTGGCCGCTCT	950
Db		848	CCATCTCTCTCATCTGTATCCAGAACCTTCAAGCAGACCTGTGATCTGGCCGCTCT	907
OY		951	TCTTCTGGGTGGGCTTTCACATTTGCTAATTCAGCCCTTAACCCCATCTCTTACAACA	1010
Db		908	TCTTCTGGGTGGGCTTTCACATTTGCTAATTCAGCCCTTAACCCCATCTCTTACAACA	967
OY		1011	TGACATCTGTGAGGAATGAGTGGAGAAATTTTTTGTGCTGTCTTGTGTTCCAGAAAAG	1070
Db		968	TGACATCTGTGAGGAATGAGTGGAGAAATTTTTTGTGCTGTCTTGTGTTCCAGAAAAG	1027
OY		1071	GAGCCATTTTAAACAGACATCTGTCAAAAAGAAATGACTTGCATTAATTTCTGGCTAA	1129
Db		1028	GAGCCATTTTAAACAGACATCTGTCAAAAAGAAATGACTTGCATTAATTTCTGGCTAA	1086

RESULT 12
AAZ49745
ID AAZ49745 standard; cDNA; 1742 BP.
XX AAZ49745;
XX
XX 02-MAY-2000 (first entry)
DE Human 14273 G-protein coupled receptor (GPCR) encoding cDNA.
XX
XX G-protein coupled receptor; GPCR, 14273 receptor; human; somatostatin;
KM cellular function/activity; galanin receptor; chemokine receptor;

KM		diagnosis; treatment; receptor-mediated disorder; screening; ss.
XX		
OS	Homo sapiens.	
XX		
FH		
FT	Key	Location/Qualifiers
FT	CDS	44..1129
FT		/tag= a
FT		/product= "Human 14273 G-Protein coupled receptor"
FT		/note= "Homologous to galanin receptor, chemokine receptor and somatostatin"
FT		/transl_except= (pos:353..355, aa:Tyr)
FT		/transl_except= (pos:446..448, aa:Asp)
FT		/transl_except= (pos:467..469, aa:Met)
FT		/transl_except= (pos:488..490, aa:Cys)
FT		/transl_except= (pos:494..496, aa:Gly)
FT		/transl_except= (pos:533..535, aa:Thr)
FT		/transl_except= (pos:740..742, aa:His)
FT		/transl_except= (pos:788..790, aa:Arg)
FT		/transl_except= (pos:791..793, aa:Ser)
FT		/transl_except= (pos:887..889, aa:Asp)
FT		/transl_except= (pos:950..952, aa:Pro)
FT		/transl_except= (pos:953..955, aa:Pro)
FT		/transl_except= (pos:965..967, aa:Ala)
FT		/transl_except= (pos:968..970, aa:Pro)
FT		458..460
FT	unSURE	/tag= b
FT		/note= "Encodes Cys"
FT		464..466
FT	unSURE	/tag= c
FT		/note= "Encodes Val"
PN	WO200000611-A2.	
XX		
PD	06-JAN-2000.	
XX		
PJ	30-JUN-1999;	99WO-US014842.
XX		
PR	30-JUN-1998;	98US-00107761.
PR	30-DEC-1998;	98US-00223538.
XX	26-FEB-1999;	99US-00261599.
PA	(MILL-) MILLENNIUM PHARM INC.	
PI	Gluckmann MA, Tsai F;	
DR	WPI: 2000-147268/13. P-P8DB: AAY44662.	
PT	Novel G-protein coupled receptor 14273 polynucleotides and polypeptides used as a target for diagnosis and treatment of receptor-mediated disorders.	
PS	Claim 3a; Fig 1; 89pp; English.	
CC	The present sequence is the cDNA encoding human G-protein coupled receptor (GPCR) 14273 protein. It is derived from human foetal brain cDNA library using expressed sequence tag (EST) that has homology to GPCR sequences. It is involved in modulation of a cellular function/activity upon binding of a ligand to the GPCR. The GPCR 14273 protein has homology to galanin receptor, chemokine receptor and somatostatin. GPCR 14273 sequences are useful as a target for diagnosis and treatment of receptor mediated disorders. The polypeptide is used for screening of drugs, agonist, antagonists and to produce antibodies. The antibodies are used for diagnosis, to assess abnormal tissue distribution or expression during development, inhibiting receptor function, tissue typing and forensic identification. The DNA sequences are used to detect genetic alterations in the gene and as a source of primers and probes	
Sequence	1742 BP; 379 A; 490 C; 439 G; 431 T; 0 U; 3 Other;	
Query Match	52.4%; Score 913; DB 3; Length 1742;	
Best Local Similarity	99.1%; Pred. No. 5.3e-311;	
Matches 1613; Conservative	0; Mismatches 14; Indels 0; Gaps 0;	

```

QY 1 TCCGGAATGATTGTAAAGACCGCTGCGGCGCCCAAGCGCCGGGAATGTCCTGAATGCGC 60
Db 1 TCGGACTAATTTCTAAGCCGCTGCGGCGCCGCAAGCCGCGGAATGTCCTGAATGCGC 60
QY 61 GCGGAGCAGCGGCGCAAGCGGCTTTGCGAGCCTGAGCAAGCAAGCCGACTTTCC 120
Db 61 GCGGAGCAGCGGCGCAAGCGGCTTTGCGAGCCTGAGCAAGCAAGCCGACTTTCC 120
QY 121 CTCTTCTTCCGAGCTCAAGGCGCAAGCCGAGCTGCTGCGCGGCTGAGACACCT 180
Db 121 CTCTTCTTCCGAGCTCAAGGCGCAAGCCGAGCTGCTGCGCGGCTGAGACACCT 180
QY 181 GCTGAGTCTATCTTTGCAAGTGTGCTGCTGAGCAAGTGTGCGCCCTGAGTGTGAGC 240
Db 181 GCTGAGTCTATCTTTGCAAGTGTGCTGCTGAGCAAGTGTGCGCCCTGAGTGTGAGC 240
QY 241 GCGCGCAGCAGCGCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 241 GCGCGCAGCAGCGCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 GCTTCTCATAGGCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GCTTCTCATAGGCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GGGGCGGCTTGGCGGCAAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 GGGGCGGCTTGGCGGCAAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 CCTCAGCCTGAGCGCGGCTGAGCCTGAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 CCTCAGCCTGAGCGCGGCTGAGCCTGAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGTGCAGGCTCTGAGCGCGCGCGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 CGTGCAGGCTCTGAGCGCGCGCGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 GGGGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 GGGGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 CGCGGACCGAGAAATTTGATGTCAGACTGATGATGATGATGATGATGATGATGATGAT 660
Db 601 CGCGGACCGAGAAATTTGATGTCAGACTGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GTGGAGATGCTCTTTTGTGATCTTTGATCTTTGATCTTTGATCTTTGATCTTTGATCT 720
Db 661 GTGGAGATGCTCTTTTGTGATCTTTGATCTTTGATCTTTGATCTTTGATCTTTGATCT 720
QY 721 TTACTCCAAAATTTTACAGATCACAAAGGCAATCAAGAAAGGCTCAAGGTAAGCTGAGC 780
Db 721 TTACTCCAAAATTTTACAGATCACAAAGGCAATCAAGAAAGGCTCAAGGTAAGCTGAGC 780
QY 781 CTACTCGAGAGCCAGCAGATCCGCGTGTCCAGCAGGACTTCCGCTCTTCCGACCT 840
Db 781 CTACTCGAGAGCCAGCAGATCCGCGTGTCCAGCAGGACTTCCGCTCTTCCGACCT 840
QY 841 CTCTCTCTCATGTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
Db 841 CTCTCTCTCATGTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
QY 901 CATCTGATCCAGAACTTCAAGCAAGACTGTGATCTGCGCGTCCCTCTTCTTCTTCTTCT 960
Db 901 CATCTGATCCAGAACTTCAAGCAAGACTGTGATCTGCGCGTCCCTCTTCTTCTTCTTCT 960
QY 961 GGTGGCTTTCATCATTTGCTAATTCAGCCCTAAACCCCATCTCTAACAATGACACTGTG 1020
Db 961 GGTGGCTTTCATCATTTGCTAATTCAGCCCTAAACCCCATCTCTAACAATGACACTGTG 1020
QY 1021 CAGGAATGATGGAAGAAATTTTGTGCTGCTTGTGTTCCAGAAAAGGAGCCATTTT 1080
Db 1021 CAGGAATGATGGAAGAAATTTTGTGCTGCTTGTGTTCCAGAAAAGGAGCCATTTT 1080

```

```

QY 1081 AACGACACATCTGTCAAAAGAAATGACTGTGATTAATTTCTGCTAATTTTCTTAT 1140
Db 1081 AACGACACATCTGTCAAAAGAAATGACTGTGATTAATTTCTGCTAATTTTCTTAT 1140
QY 1141 AGCCGAGTTTCTCAACCTGCGGAGCTGTGCGCATGCTTTTAAACAGAGTTCAATTCAGT 1200
Db 1141 AGCCGAGTTTCTCAACCTGCGGAGCTGTGCGCATGCTTTTAAACAGAGTTCAATTCAGT 1200
QY 1201 ACCCTCATCAGTCAACCCCTGTTTAAAGAAATGAACCTTAGCAATAGACATCCAGC 1260
Db 1201 ACCCTCATCAGTCAACCCCTGTTTAAAGAAATGAACCTTAGCAATAGACATCCAGC 1260
QY 1261 GTGCGTAATTAAGGCGTATCAACCAAGTTCTAATATTTTCCCTTATPAAAGATTT 1320
Db 1261 GTGCGTAATTAAGGCGTATCAACCAAGTTCTAATATTTTCCCTTATPAAAGATTT 1320
QY 1321 GTTGGCCAGGTGCAAGTGTTCATGCTCTGTAATCCGACAGTTTGGAGGCTGAGTGGGT 1380
Db 1321 GTTGGCCAGGTGCAAGTGTTCATGCTCTGTAATCCGACAGTTTGGAGGCTGAGTGGGT 1380
QY 1381 GGATCACTGAGGTCAAGAGTTGAGACCAACCTGACCAACATGATGAGACCCCGTCTC 1440
Db 1381 GGATCACTGAGGTCAAGAGTTGAGACCAACCTGACCAACATGATGAGACCCCGTCTC 1440
QY 1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1500
Db 1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1500
QY 1501 CTTGGAGGCTGAACCGAGGAATCTTCTTAACCTGGAGGCGAGAGTTTCAAGTGAACCG 1560
Db 1501 CTTGGAGGCTGAACCGAGGAATCTTCTTAACCTGGAGGCGAGAGTTTCAAGTGAACCG 1560
QY 1561 AGATCGTGCACTTCCCAACGAGGCAACAGAGTGAATCTCATCTTTAAAAAAA 1620
Db 1561 AGATCGTGCACTTCCCAACGAGGCAACAGAGTGAATCTCATCTTTAAAAAAA 1620
QY 1621 AAAAAA 1627
Db 1621 AAAAAA 1627

RESULT 13
ABA21432
ID ABA21432 standard; DNA; 746 BP.
XX
AC ABA21432:
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 13763.
XX
KW Human; nootropic; neuroprotective; cytoskeletal; dermatological; vinorelbine;
KW immunosuppressive; anti-infective; anti-inflamatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antispasmodic; antianemic; antidiabetic; cancer;
KW antileukemic; hepatotropic; cerebroprotective; anti-inflammatory;
KW antileukemic; antidiabetic; antileukemic; antileukemic; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN MO200159063-A2.
XX
PD 16-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.

```

PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231433P.
PR 08-SEP-2000; 2000US-0231434P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.

PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251566P.
PR 08-DEC-2000; 2000US-0251688P.
PR 08-DEC-2000; 2000US-0251699P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
PR Rosen CA, Barash SC, Ruben SM,
XX WPI; 2001-541565/60.
XX
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX
XX Disclosure; SEQ ID NO 13763; 1701bp + Sequence Listing; English.
PS
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are

isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIP0 at ftp.wip0.int/pub/published_pct_sequences
 XX
 SQ Sequence 746 BP; 224 A; 149 C; 164 G; 209 T; 0 U; 0 Other;
 Query Match 42.8%; Score 746; DB 5; Length 746;
 Best Local Similarity 100.0%; Pred. No. 2.4e-252;
 Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 976 TGCCTAATTCAGCCCTTAACCCCATCTCTCAACATGACATGCGCGAATAGTGGAA 1035
 DB 1 TGCCTAATTCAGCCCTTAACCCCATCTCTCAACATGACATGCGCGAATAGTGGAA 60
 QY 1036 GAAATTTTTCGCTGCTTGGTTCCGAGAAAGGAGCCATTTTACAGACATCTGT 1095
 DB 61 GAAATTTTTCGCTGCTTGGTTCCGAGAAAGGAGCCATTTTACAGACATCTGT 120
 QY 1096 CAAAGAAATGACTGTGATTAATTTCTGCTAATTTTCTTTATAGCCGAGTTTCTCAC 1155
 DB 121 CAAAGAAATGACTGTGATTAATTTCTGCTAATTTTCTTTATAGCCGAGTTTCTCAC 180
 QY 1156 ACCTGGGAGCTGTGGATGCTTTTAAACAGAGTTCAATTTCCAGTACCTCCCAACAGTGC 1215
 DB 181 ACCTGGGAGCTGTGGATGCTTTTAAACAGAGTTCAATTTCCAGTACCTCCCAACAGTGC 240
 QY 1216 ACCCTGCTTAAAGAAATGAACCTATGCAATAGACATCCACAGCGTGGTAAATTAAAG 1275
 DB 241 ACCCTGCTTAAAGAAATGAACCTATGCAATAGACATCCACAGCGTGGTAAATTAAAG 300
 QY 1276 GGTGATCACCAGATTGATATATTTTCCCTTAAAGAGATTGTTGGCCAGTGCAG 1335
 DB 301 GGTGATCACCAGATTGATATATTTTCCCTTAAAGAGATTGTTGGCCAGTGCAG 360
 QY 1336 TGGTTCAATGCTGTAAATCCACAGAGTTGGAGGCTGAGTGGTGGATCACTGAGGTC 1395
 DB 361 TGGTTCAATGCTGTAAATCCACAGAGTTGGAGGCTGAGTGGTGGATCACTGAGGTC 420
 QY 1396 AGGAGTTGAGACCAACCTGACCAACATGGAGACCCCGTCTCTAATAAATAAATAA 1455
 DB 421 AGGAGTTGAGACCAACCTGACCAACATGGAGACCCCGTCTCTAATAAATAAATAA 480
 QY 1456 AAAAAATTAAGCTGGAGTGTGTGGGCACTGTAAATCTAGTACTTGGAGGCTGAGC 1515
 DB 481 AAAAAATTAAGCTGGAGTGTGTGGGCACTGTAAATCTAGTACTTGGAGGCTGAGC 540
 QY 1516 CAGAGAAATCTCTGAACCTGGAGGAGAGAGTGTGAGTGAAGCCGAATGCTGCATTGC 1575
 DB 541 CAGAGAAATCTCTGAACCTGGAGGAGAGAGTGTGAGTGAAGCCGAATGCTGCATTGC 600
 QY 1576 ACTCCACACGAGGAGCAAGAGTGAATCTGCATCTTAAAAAATAAATAAATAAATTGT 1635
 DB 601 ACTCCACACGAGGAGCAAGAGTGAATCTGCATCTTAAAAAATAAATAAATAAATTGT 660
 QY 1636 TATGGTTCCTTTTAAATGTGAATTTTAAATGTGTGTTGTAATATGATCAAAATTTAATA 1695
 DB 661 TATGGTTCCTTTTAAATGTGAATTTTAAATGTGTGTTGTAATATGATCAAAATTTAATA 720
 QY 1696 AATATTTTATATGATGCTTACAGCA 1721
 DB 721 AATATTTTATATGATGCTTACAGCA 746

RESULT 14
 ACA61000
 ID ACA61000 standard; DNA; 632 BP.
 XX
 AC ACA61000;
 XX
 DT 04-JUL-2003 (first entry)
 XX
 DE Novel human G protein coupled receptor HGPBMY18 3' UTR.
 XX
 XX Human; G protein coupled receptor; GPCR; HGPBMY18; cytosolic;
 KW neurotropic; neuroprotective; antiparkinsonian; antidiabetic;
 KW antidepressant; hypotensive; tranquilizer; nephrotoxic;
 KW cell growth disorder; neurological condition; pituitary gland disorder;
 KW colon disorder; breast disorder; lung disorder; prostate disorder;
 KW Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;
 KW colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;
 KW sleeplessness; hypertension; anxiety; stress; renal failure; 3'UTR;
 KW untranslated region; ds.
 XX
 OS Homo sapiens.
 XX
 PN US2003022186-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 14-NOV-2001; 2001US-00992331.
 XX
 PR 14-NOV-2000; 2000US-0248483P.
 PR 16-JUN-2001; 2001US-0261782P.
 PR 27-JUL-2001; 2001US-0308540P.
 XX
 PA (FEDER/) FEDER J N.
 PA (MINT/) MINTIER G.
 PA (PAMA/) RAMANATHAN C S.
 XX
 PI Feder JN, Mintier G, Ramanathan CS;
 XX
 DR WP1; 2003-416985/39.
 XX
 PT New isolated human GPCR, HGPBMY18, polynucleotide and polypeptide,
 PT useful for treating, diagnosing and preventing disorders such as cancer,
 PT neurological conditions, and diseases of the pituitary gland, colon,
 PT breast and lungs.
 XX
 PS Disclosure; Fig 4; 66pp; English.
 XX
 CC The invention describes a new isolated polynucleotide encoding a human G-
 CC protein coupled receptor (GPCR) or its functional fragment. The methods
 CC and compositions of the present invention are useful for treating,
 CC diagnosing, preventing and screening disorders associated with aberrant
 CC cell growth, neurological conditions, and diseases related to the
 CC pituitary gland, colon, breast, lungs and prostate. These disorders
 CC include Alzheimer's, Parkinson, diabetes, dwarfism, colour blindness,
 CC retinal pigmentosa, asthma, depression, schizophrenia, sleeplessness,
 CC hypertension, anxiety, stress and renal failure. This sequence represents
 CC novel human G protein coupled receptor HGPBMY18 3' untranslated region
 CC (UTR)
 XX
 SQ Sequence 632 BP; 219 A; 113 C; 134 G; 166 T; 0 U; 0 Other;
 Query Match 33.6%; Score 586; DB 7; Length 632;
 Best Local Similarity 100.0%; Pred. No. 3.5e-196;
 Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1130 TTTTCTTTTAAAGCCGAGTTTCTCAACCTGGAGAGCTGTGGATGCTTTTAAACAGAGT 1189
 DB 1 TTTTCTTTTAAAGCCGAGTTTCTCAACCTGGAGAGCTGTGGATGCTTTTAAACAGAGT 60
 QY 1190 TCATTTCCAGTACCTTCATCAGTGCACCTGCTTTTAAAGAAATGAACCTATGCAATAG 1249
 DB 61 TCATTTCCAGTACCTTCATCAGTGCACCTGCTTTTAAAGAAATGAACCTATGCAATAG 120

Db 301 GCTCTCATGAGCGCTATCCCTCTGCTGCGCGGCTGAGCTGAGCGCTGCTGCT 360
 Qy 361 GGGGCGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 GGGGCGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Qy 421 CTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 421 CTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Qy 481 CGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 481 CGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Qy 541 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 541 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Qy 601 CGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 601 CGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Qy 661 GTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 GTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Qy 721 TTAATCCAAATTTTACAGATCAAGAGGATCAAGAGGCTCAAGAGGCTCAAGAGGCT 780
 Db 721 TTAATCCAAATTTTACAGATCAAGAGGATCAAGAGGCTCAAGAGGCTCAAGAGGCT 780
 Qy 781 CTACTCGGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGG 840
 Db 781 CTACTCGGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGG 840
 Qy 841 CTTCCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 841 CTTCCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Qy 901 CATCTGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCA 960
 Db 901 CATCTGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCAAGAGGATCA 960
 Qy 961 GGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 961 GGTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Qy 1021 CAGGAATGATGGAAGAAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1021 CAGGAATGATGGAAGAAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Qy 1081 AACGAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 Db 1081 AACGAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 Qy 1141 AGCGGAGCTTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 Db 1141 AGCGGAGCTTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 Qy 1201 ACCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 1201 ACCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Qy 1261 GTGGGATTAATTAAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Db 1261 GTGGGATTAATTAAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Qy 1321 GTGGGATTAATTAAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Db 1321 GTGGGATTAATTAAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Qy 1381 GGAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Db 1381 GGAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440

Qy 1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
 Db 1441 TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
 Qy 1501 CTTGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 1501 CTTGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Qy 1561 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 1561 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Qy 1621 AAAAAAAGATTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 1621 AAAAAAAGATTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Qy 1681 TGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
 Db 1681 TGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
 Qy 1741 CGG 1743
 Db 1741 CGG 1743
 RESULT 3
 US-09-345-882-1/c
 : Sequence 1, Application US/09345882
 : Patent No. 639373
 : GENERAL INFORMATION:
 : APPLICANT: Bouquelieret, Lydie
 : TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
 : FILE REFERENCE: GENSET.031A
 : CURRENT APPLICATION NUMBER: US/09/345,882
 : PRIOR FILING DATE: 1999-06-30
 : PRIOR APPLICATION NUMBER: US 60/091,315
 : PRIOR FILING DATE: 1998-06-30
 : NUMBER OF SEQ ID NOS: 140
 : SOFTWARE: Patent.pm
 : SEQ ID NO 1
 : LENGTH: 162450
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: allele
 : LOCATION: 72794 : polymorphic base A or G
 : OTHER INFORMATION: 5-124-273 : polymorphic base A or G
 : FEATURE:
 : NAME/KEY: allele
 : LOCATION: 88073 : polymorphic base A or C
 : OTHER INFORMATION: 5-127-261 : polymorphic base A or C
 : FEATURE:
 : NAME/KEY: allele
 : LOCATION: 90842 : polymorphic base A or G
 : OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
 : FEATURE:
 : NAME/KEY: allele
 : LOCATION: 93714 : polymorphic base deletion of GT
 : OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
 : FEATURE:
 : NAME/KEY: allele
 : LOCATION: 97122 : polymorphic base G or T
 : OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
 : FEATURE:
 : NAME/KEY: allele
 : LOCATION: 97152 : polymorphic base deletion of T
 : OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
 : FEATURE:
 : NAME/KEY: allele

LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096

OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59

```
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match 3.3%; Score 58; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 7.3e-14;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 CTTGAACCTGGAGGCGAGGTTGCACTGAGCCGAGATCGTCCATTGCACTCCACC 1584
DB 8387 CTTGAACCTGGAGGCGAGGTTGCACTGAGCCGAGATCGTCCATTGCACTCCACC 8330

RESULT 4
US-09-482-273-79
Sequence 79, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 79
LENGTH: 2191
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1327)
OTHER INFORMATION: n equals a,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1334)
OTHER INFORMATION: n equals a,c,g, or c
US-09-482-273-79

Query Match 2.9%; Score 50; DB 4; Length 2191;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGGAGGCTGAGGTGGTGATCAGCTGAGGTGAGAGTTGAGACCA 1410
DB 1929 TTGGAGGCTGAGGTGGTGATCAGCTGAGGTGAGAGTTGAGACCA 1978

RESULT 5
US-09-564-805-28/c
Sequence 28, Application US/09564805
Patent No. 6333403
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
```

```
TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
FILE REFERENCE: 2318-258
CURRENT APPLICATION NUMBER: US/09/564,805
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US 60/107,468
PRIOR FILING DATE: 1998-11-06
PRIOR APPLICATION NUMBER: 09/434,382
PRIOR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 240
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 28
LENGTH: 2664
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (910)..(13104)
OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
OTHER INFORMATION: exon 8: 8186-8244; exon 9: 12878-12936; exon 10:
NAME/KEY: misc feature
LOCATION: (13756)..(22917)
OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
NAME/KEY: misc feature
LOCATION: (23045)..(26452)
OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
OTHER INFORMATION: signal: 26447-26452
NAME/KEY: variation
LOCATION: (826)..(23879)
OTHER INFORMATION: s at positions 826 and 23180 is G or C; Y at
OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
OTHER INFORMATION: 18 C or T; n at position 13128 is C or G; T at
OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-564-805-28

Query Match 2.9%; Score 50; DB 4; Length 2664;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGGAGGCTGAGGTGGTGATCAGCTGAGGTGAGAGTTGAGACCA 1410
DB 17609 TTGGAGGCTGAGGTGGTGATCAGCTGAGGTGAGAGTTGAGACCA 17560

RESULT 6
US-09-820-924-3
Sequence 3, Application US/09820924
Patent No. 6555351
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M. et al
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001213
CURRENT APPLICATION NUMBER: US/09/820,924
CURRENT FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 39982
TYPE: DNA
ORGANISM: Human
US-09-820-924-3

Query Match 2.9%; Score 50; DB 4; Length 39982;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1361 TTTGGAGGCTGAGTGGTGGATCAGCTGAGTGGAGGTTGGAGCA 1410
DB 22437 TTTGGAGGCTGAGTGGTGGATCAGCTGAGTGGAGGTTGGAGCA 22486

RESULT 7

US-09-369-247-46
; Sequence 46: Application US/09369247
; Patent No. 6569992
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 44 Human Secreted Proteins
; FILE REFERENCE: PZ024P1
; CURRENT APPLICATION NUMBER: US/09/369,247
; CURRENT FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: 60/074,118
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,157
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,137
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,341
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,141
; EARLIER FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 2094
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-369-247-46

Query Match

Best Local Similarity 100.0%; Score 49; DB 4; Length 2094;
Pred. No. 3,7e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 CAGGAGATCTCTGAGCTGGAGGAGGAGGTTGAGTGGAGGAT 1564
DB 1975 CAGGAGATCTCTGAGCTGGAGGAGGAGGTTGAGTGGAGGAT 2023

RESULT 8

US-09-621-976-14833
; Sequence 14833: Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent .pm
; SEQ ID NO 14833
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14833

Query Match

Best Local Similarity 100.0%; Score 48; DB 4; Length 405;
Pred. No. 1.1e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1345 CCTGTAATCCAGCAGTTTGGAGGCTGAGTGGGATCAGCTGAG 1392
DB 191 CCTGTAATCCAGCAGTTTGGAGGCTGAGTGGGATCAGCTGAG 238

RESULT 9

US-09-621-976-15040
; Sequence 15040: Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent .pm
; SEQ ID NO 15040
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15040

Query Match 2.8%; Score 48; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1345 CCTGTAATCCAGCAGTTTGGAGGCTGAGTGGGATCAGCTGAG 1392
DB 191 CCTGTAATCCAGCAGTTTGGAGGCTGAGTGGGATCAGCTGAG 238

RESULT 10

US-09-435-739-42/c
; Sequence 42: Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Pecker, Iris
; APPLICANT: Viodevsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435,739
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 44848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-435-739-42

Query Match 2.8%; Score 48; DB 4; Length 44848;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGGAGGCTGAGTGGTGGATCAGCTGAGTGGAGGTTGGAGAC 1408
DB 19123 TTTGGAGGCTGAGTGGTGGATCAGCTGAGTGGAGGTTGGAGAC 19076

RESULT 11

US-09-497-855A-38/c
; Sequence 38: Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54

QY 1361 TTTGGAGGCTGAGTGGTGGATCAGCTGAGTGGAGGTTGGAGAC 1408
DB 19123 TTTGGAGGCTGAGTGGTGGATCAGCTGAGTGGAGGTTGGAGAC 19076

SOFTWARE: PatentIn version 3.0
SEQ ID NO 38779
LENGTH: 128779
TYPE: DNA
ORGANISM: Homo sapiens;
US-09-497-855A-38

Query Match 2.7%; Score 47; DB 4; Length 128779;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1520 AGAATCTCTGAACCTGGAGGAGAGGAGTTCAGTGCAGATCG 1566
Db 94386 AGAATCTCTGAACCTGGAGGAGAGGAGTTCAGTGCAGATCG 94340

RESULT 12
US-08-451-777A-7/C
Sequence 7, Application US/08451777A
Patent No. 5789223
GENERAL INFORMATION:
APPLICANT: Bergsma, Derk J.
APPLICANT: Stambolian, Dwight
TITLE OF INVENTION: Human Galactokinase Gene
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESSEE: Intellectual Property
STREET: 709 Swedeland Road/UW2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,777A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Eagle, Alissa M.
REGISTRATION NUMBER: 37,126
REFERENCE/DOCKET NUMBER: P50268-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5090
TELEFAX: 610-270-5364
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-451-777A-7

Query Match 2.6%; Score 45; DB 1; Length 7676;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1527 CTTGAACCTGGAGGAGAGGAGTTCAGTGCAGATCGTGC 1571
Db 4083 CTTGAACCTGGAGGAGGAGGAGTTCAGTGCAGATCGTGC 4039

RESULT 13
US-08-451-778A-7/C
Sequence 7, Application US/08451778A

Patent No. 5830649
GENERAL INFORMATION:
APPLICANT: Bergsma, Derk J.
APPLICANT: Stambolian, Dwight
TITLE OF INVENTION: Human Galactokinase Gene
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESSEE: Intellectual Property
STREET: 709 Swedeland Road/UW2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,778A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Eagle, Alissa M.
REGISTRATION NUMBER: 37,126
REFERENCE/DOCKET NUMBER: P50268-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5364
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-451-778A-7

Query Match 2.6%; Score 45; DB 2; Length 7676;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1527 CTTGAACCTGGAGGAGAGGAGTTCAGTGCAGATCGTGC 1571
Db 4083 CTTGAACCTGGAGGAGGAGGAGTTCAGTGCAGATCGTGC 4039

RESULT 14
US-08-998-208-7/C
Sequence 7, Application US/08998208
Patent No. 5880105
GENERAL INFORMATION:
APPLICANT: Bergsma, Derk J.
APPLICANT: Stambolian, Dwight
TITLE OF INVENTION: Human Galactokinase Gene
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESSEE: Intellectual Property
STREET: 709 Swedeland Road/UW2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/998,208
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA: 08/451,777
;; APPLICATION NUMBER: 08/451,777
;; FILING DATE: 26-MAY-1995
;; APPLICATION NUMBER: PCT/US94/10825
;; FILING DATE: 23-SEP-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Eagle, Alisha M.
;; REGISTRATION NUMBER: 37,126
;; REFERENCE//DOCKET NUMBER: P50268-1B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-270-5364
;; TELEFAX: 610-270-5090
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7676 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-998-208-7

Query Match 2.6%; Score 45; DB 2; Length 7676;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 CTTGAACCTGGAGGAGGAGGTTGCAGTGAGCCGAGATCGTCCA 1571
DB 4083 CTTGAACCTGGAGGAGGAGGTTGCAGTGAGCCGAGATCGTCCA 4039

RESULT 15
PCT-US95-06743-7/c
; Sequence 7, Application PC/TUS9506743
; GENERAL INFORMATION:
; APPLICANT: Bergsma, Derk J.
; APPLICANT: Stambolian, Dwight
; TITLE OF INVENTION: Human Galactokinase Gene
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; ADDRESSEE: Intellectual Property
; STREET: 709 Swedeland Road/TW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10825
; FILING DATE: 23-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Suton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE//DOCKET NUMBER: P50268-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5024
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7676 base pairs
; TYPE: nucleic acid

;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; PCT-US95-06743-7

Query Match 2.6%; Score 45; DB 5; Length 7676;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 CTTGAACCTGGAGGAGGAGGTTGCAGTGAGCCGAGATCGTCCA 1571
DB 4083 CTTGAACCTGGAGGAGGAGGTTGCAGTGAGCCGAGATCGTCCA 4039

Search completed: October 1, 2004, 03:56:15
Job time : 141.258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 30, 2004, 21:52:37 ; Search time 817.411 Seconds
(without alignments)
10810.058 Million cell updates/sec

Title: US-10-077-698-2

Perfect score: 1743

Sequence: 1 tccggactgtcttagacg.....aaaaaaaaaaggcg 1743

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3340653 seqs, 2534783454 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1743	100.0	1743	14	US-10-086-181-1
2	1743	100.0	1743	15	US-10-077-698-2
3	1743	100.0	1743	15	US-10-171-027-2
4	1743	100.0	1743	15	US-10-075-987-2
5	1109	63.6	1160	14	US-10-015-498-1
6	1086	62.3	1086	14	US-10-086-181-3
7	1086	62.3	1086	15	US-10-083-168-11
8	1035	59.4	1086	15	US-09-892-331-1
9	1035	59.4	1086	15	US-10-083-168-78
10	1035	59.4	1086	15	US-10-262-313-1
11	1035	59.4	1086	15	US-10-768-878-1
12	984	56.5	1086	17	US-10-083-168-80
13	926	53.1	1086	9	US-09-995-325-7
14	926	53.1	1086	10	US-09-995-325-7

15	586	33.6	632	10	US-09-992-331-4	Sequence 4, Appl1
16	586	33.6	632	15	US-10-262-313-4	Sequence 4, Appl1
17	586	33.6	632	17	US-10-768-878-4	Sequence 4, Appl1
18	559	32.1	1066	15	US-10-017-161-1809	Sequence 1809, Ap
19	559	32.1	1066	16	US-10-292-798-1465	Sequence 1465, Ap
20	550	31.6	3173	15	US-10-116-252-5	Sequence 5, Appl1
21	391	22.4	1737	15	US-10-225-5678-681	Sequence 681, App
22	337	19.3	368	13	US-10-276-774-265	Sequence 265, App
23	326	18.7	1104	10	US-09-791-932-10	Sequence 10, Appl
24	298	17.1	760	13	US-10-076-553-771	Sequence 771, App
25	198	11.4	300	13	US-10-076-555-45	Sequence 45, Appl
26	143	8.2	300	9	US-10-076-555-46	Sequence 46, Appl
27	131	7.5	638	9	US-09-828-644-58	Sequence 58, Appl
28	111	6.4	426	10	US-09-791-932-33	Sequence 33, Appl
29	66	3.8	25871	9	US-09-768-7434-5	Sequence 5, Appl1
30	65	3.7	178896	17	US-10-450-826-53	Sequence 53, Appl
31	58	3.3	2378	13	US-10-027-632-101748	Sequence 101748,
32	58	3.3	2378	16	US-10-027-632-101748	Sequence 101748,
33	58	3.3	162450	15	US-10-071-179-1	Sequence 1, Appl1
34	58	3.3	162450	15	US-10-126-704-1	Sequence 1, Appl1
35	57	3.3	80	10	US-09-392-331-5	Sequence 5, Appl1
36	57	3.3	80	15	US-10-262-313-5	Sequence 5, Appl1
37	57	3.3	80	15	US-10-768-878-5	Sequence 5, Appl1
38	57	3.3	19553	9	US-09-764-847-1425	Sequence 1425, Ap
39	57	3.3	19553	15	US-10-092-154-1425	Sequence 1425, Ap
40	55	3.2	3319	17	US-10-474-495-143	Sequence 143, App
41	55	3.2	3287	13	US-10-027-632-113282	Sequence 113282,
42	55	3.2	3287	13	US-10-027-632-113283	Sequence 113283,
43	55	3.2	3287	13	US-10-027-632-113284	Sequence 113284,
44	55	3.2	3287	16	US-10-027-632-113282	Sequence 113282,
45	55	3.2	3287	16	US-10-027-632-113283	Sequence 113283,

ALIGNMENTS

RESULT 1

US-10-086-181-1

Sequence 1, Application US/10086181

Publication No. US20020177151A1

GENERAL INFORMATION:

APPLICANT: GIMENO, Ruth

TITLE OF INVENTION: METHODS FOR THE TREATMENT OF METABOLIC DISORDERS, INCLUDING OBESITY AND DIABETES

FILE REFERENCE: NMI-220

CURRENT APPLICATION NUMBER: US/10/086.181

CURRENT FILING DATE: 2002-02-26

PRIOR APPLICATION NUMBER: 60/271,655

PRIOR FILING DATE: 2001-02-26

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1743

TYPE: DNA

ORGANISM: Homo sapiens

FEATURES:

NAME/KEY: CDS

LOCATION: (44)...(1129)

US-10-086-181-1

Query Match	100.0%	Score 1743;	DB 14;	Length 1743;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1743;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	TCCGAGCTAGTTCTAGACGCGCGGCGCGCGGAGATGCTCCCTGATGCGC	60	
DB	1	TCCGAGCTAGTTCTAGACGCGCGCGCGCGCGGAGATGCTCCCTGATGCGC	60	
QY	61	GCGGAGCGCGCGCGCGCGCGCGCTTTCGCGACCTTGAGAGCAACCGACCGCTTTC	120	
DB	61	GCGGAGCGCGCGCGCGCGCGCGCTTTCGCGACCTTGAGAGCAACCGACCGCTTTC	120	
QY	121	CTTCTTCTCGAGAGTCAAGGCGACCGCGCTGTCTCTGCGCGCGGTGAGACACCGT	180	

```
Db 121 |||||CTTCTTCGCGAGTGAAGGCGACACCGGCTGTGCTGGCCGGGTGGAGAACCGT 180
Qy 181 |||||GCTGTGTCTATCTTTGCACTGTGCTGTGGGCAAGTGTGGCCCTGTGTGTGGC 240
Db 181 |||||GCTGTGTCTATCTTTGCACTGTGCTGTGGGCAAGTGTGGCCCTGTGTGTGGC 240
Qy 241 |||||GGCGCGACGACGCGCGCGCGGACGCTGTGCTGTGTACTGAACCTCTTCTGCGGAGCT 300
Db 241 |||||GGCGCGACGACGCGCGCGCGGACGCTGTGCTGTGTACTGAACCTCTTCTGCGGAGCT 300
Qy 301 |||||GCTCTTCATCAGCGCTATCCCTCTGTGTGTGCTGTGCGGTGCGCTGGAAGTGGCTGCT 360
Db 301 |||||GCTCTTCATCAGCGCTATCCCTCTGTGTGTGCTGTGCGGTGCGCTGGAAGTGGCTGCT 360
Qy 361 |||||GGGCGCCCGTGTGCTGCGCACTGTCTTCTACGTGATGACCTTGAGCGGCGTCAACAT 420
Db 361 |||||GGGCGCCCGTGTGCTGCGCACTGTCTTCTACGTGATGACCTTGAGCGGCGTCAACAT 420
Qy 421 |||||CCTCAAGCTGGCGCGGCTGAGCGCTGAGCGCATGTGTGATCGTGAACCTGACGCGCG 480
Db 421 |||||CCTCAAGCTGGCGCGGCTGAGCGCTGAGCGCATGTGTGATCGTGAACCTGACGCGCG 480
Qy 481 |||||CGTGTGGGGTCTGTGGGCGCGGCGCGGAGTGTGCTGTGCGCTCATCTGTGGGCTATTC 540
Db 481 |||||CGTGTGGGGTCTGTGGGCGCGGCGCGGAGTGTGCTGTGCGCTCATCTGTGGGCTATTC 540
Qy 541 |||||GGGGGTGTGCGCTGTGCTGTGCGCTTCTTCAAGTGTGCTGCGCAAGGCTGCGCGG 600
Db 541 |||||GGGGGTGTGCGCTGTGCTGTGCGCTTCTTCAAGTGTGCTGCGCAAGGCTGCGCGG 600
Qy 601 |||||CGCGGACGAGAAATTTGATTTGCACTGATTTGGCCCACTTCCTGGAAGATCTC 660
Db 601 |||||CGCGGACGAGAAATTTGATTTGCACTGATTTGGCCCACTTCCTGGAAGATCTC 660
Qy 661 |||||GTGGGATGTCTCTTTGTGTAATTTGAACTTCTGTGTGCGAGGATGTGTGATCAG 720
Db 661 |||||GTGGGATGTCTCTTTGTGTAATTTGAACTTCTGTGTGCGAGGATGTGTGATCAG 720
Qy 721 |||||TTTACTCCAAATTTTAAAGATCAACAAAGGATCAAGAAAGGCTCAAGCTGGTGGC 780
Db 721 |||||TTTACTCCAAATTTTAAAGATCAACAAAGGATCAAGAAAGGCTCAAGCTGGTGGC 780
Qy 781 |||||CTACTCGGAGGACCAAGATCGCGGTGCGGAGGAGTCTTCCGACCT 840
Db 781 |||||CTACTCGGAGGACCAAGATCGCGGTGCGGAGGAGTCTTCCGACCT 840
Qy 841 |||||CTTCTCTCATGTGTCTCTTCTTCAATGTGTGAGGCCCATCATCATCATCTCT 900
Db 841 |||||CTTCTCTCATGTGTCTCTTCTTCAATGTGTGAGGCCCATCATCATCATCTCT 900
Qy 901 |||||CATCTGTATCCAAATTTGAAAGGACCTGTGTATGTGCGCGCTCTTCTTCTGTGGT 960
Db 901 |||||CATCTGTATCCAAATTTGAAAGGACCTGTGTATGTGCGCGCTCTTCTTCTGTGGT 960
Qy 961 |||||GGTGGCCTTCACTTTGCTAATTCAGGCTTAAACCCCATCTCTACAAACATGACATGTG 1020
Db 961 |||||GGTGGCCTTCACTTTGCTAATTCAGGCTTAAACCCCATCTCTACAAACATGACATGTG 1020
Qy 1021 |||||CAGGAATGAGTGAAGAAATTTTGTGCTGTGCTGTGCTGCAAGAAAGGAGGCTATTT 1080
Db 1021 |||||CAGGAATGAGTGAAGAAATTTTGTGCTGTGCTGTGCTGCAAGAAAGGAGGCTATTT 1080
Qy 1081 |||||AACAGAACATCTGTCAAAAGAAATGACTGTGATATTTTGTGGCTAATTTTCTTTAT 1140
Db 1081 |||||AACAGAACATCTGTCAAAAGAAATGACTGTGATATTTTGTGGCTAATTTTCTTTAT 1140
Qy 1141 |||||AGCGGATTTTCTCACTGTGGGAGCTGTGGCATGCTTTTAAAGAGATTCATTTCCAGT 1200
Db 1141 |||||AGCGGATTTTCTCACTGTGGGAGCTGTGGCATGCTTTTAAAGAGATTCATTTCCAGT 1200
Qy 1201 |||||ACCTTCATGATGACCTGTCTTTAAGAAATGAACTATGCAATGACATCCAGC 1260
Db 1201 |||||ACCTTCATGATGACCTGTCTTTAAGAAATGAACTATGCAATGACATCCAGC 1260
```

```
Db 1201 |||||ACCTTCATGATGACCTGTCTTTAAGAAATGAACTATGCAATGACATCCAGC 1260
Qy 1261 |||||GTGGTAAATTAAGGGGTGATGCAAGATTGATTAATTTTCTTATTAAGATTT 1320
Db 1261 |||||GTGGTAAATTAAGGGGTGATGCAAGATTGATTAATTTTCTTATTAAGATTT 1320
Qy 1321 |||||GTTGGCAGGTGAGTGTGCTGATGCTGTATGCCAGCACTTTGGAGGCTGAGTGGT 1380
Db 1321 |||||GTTGGCAGGTGAGTGTGCTGATGCTGTATGCCAGCACTTTGGAGGCTGAGTGGT 1380
Qy 1381 |||||GATTCACCTGAGGTGAGAGTGTGAGACCACTGACCAACATGTGTAGACCCCGCTC 1440
Db 1381 |||||GATTCACCTGAGGTGAGAGTGTGAGACCACTGACCAACATGTGTAGACCCCGCTC 1440
Qy 1441 |||||TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1500
Db 1441 |||||TACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1500
Qy 1501 |||||CTTGGAGGCTGAACGAGAAATCTTTGAACTGTGGAGGCAAGGTTGCAAGTGGC 1560
Db 1501 |||||CTTGGAGGCTGAACGAGAAATCTTTGAACTGTGGAGGCAAGGTTGCAAGTGGC 1560
Qy 1561 |||||AGATCGTCCATTCACCTCCACCAAGGCAACAGAGTGAATCTCTTAATTAATTAAT 1620
Db 1561 |||||AGATCGTCCATTCACCTCCACCAAGGCAACAGAGTGAATCTCTTAATTAATTAAT 1620
Qy 1621 |||||AAAAAAGATTTGTTATGGGCTCTTTTAAATGTGAATTTTATGTGTGTGATA 1680
Db 1621 |||||AAAAAAGATTTGTTATGGGCTCTTTTAAATGTGAATTTTATGTGTGTGATA 1680
Qy 1681 |||||TGATCAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1740
Db 1681 |||||TGATCAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1740
Qy 1741 |||||CGG 1743
Db 1741 |||||CGG 1743
```

```
RESULT 2
US-10-077-698-2
; Sequence 2, Application US/10077698
; Publication No. US20030008350A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14273 Receptor, A No. US20030008350A1 G-Protein Coupled Receptor
; FILE REFERENCE: 5800-4B, 035800/177086
; CURRENT APPLICATION NUMBER: US/10/077,698
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 09/261,599
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/107,761
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/223,538
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-077-698-2
```

```
Query Match 100.0%; Score 1743; DB 15; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 |||||TCCGACTAGATTCTAGACCGCTCGGCGCGGAGGATGTCCTCTGAATGCGC 60
Db 1 |||||TCCGACTAGATTCTAGACCGCTCGGCGCGGAGGATGTCCTCTGAATGCGC 60
Qy 61 |||||GGGCGACGCGGCGACGCGGCTTGTGCGACCTGTGAGCAAGCAACGCAACCGGCTTTCC 120
Db 61 |||||GGGCGACGCGGCGACGCGGCTTGTGCGACCTGTGAGCAAGCAACGCAACCGGCTTTCC 120
```



```

QY 351 CCGGCTGCTGGGCCCCGCTTGGCTGCACTGCTCTTCTAAGTATGACCTTGAGGGCA 410
DB 308 CCGGCTGCTGGGCCCCGCTTGGCTGCACTGCTCTTCTAAGTATGACCTTGAGGGCA 367
QY 411 GCGTCAACCATCTCCACGCTGGCGCGGTGAGCTGAGCCATGATGTGATGATGAC 470
DB 368 GCGTCAACCATCTCCACGCTGGCGCGGTGAGCTGAGCCATGATGTGATGATGAC 427
QY 471 TCGAGCGCGCGGTGAGCTGAGCGCGGTGAGCTGAGCGCGGTGAGCTGAGCGCTCATCT 530
DB 428 TCGAGCGCGCGGTGAGCTGAGCGCGGTGAGCTGAGCGCGGTGAGCTGAGCGCTCATCT 487
QY 531 GGGGCTATTTGGGCGGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 590
DB 488 GGGGCTATTTGGGCGGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 547
QY 591 GGGTCCCCGGCGCGGCAACGAGAAATTTGATTTGACATGATTTGGCCCACTTCCTG 650
DB 548 GGGTCCCCGGCGCGGCAACGAGAAATTTGATTTGACATGATTTGGCCCACTTCCTG 607
QY 651 GAGAGATCTGCTGGGATGCTCTTTGTTGATCTTTGATCTTTGATCTTTGATCTTTGATCT 710
DB 608 GAGAGATCTGCTGGGATGCTCTTTGTTGATCTTTGATCTTTGATCTTTGATCTTTGATCT 667
QY 711 TTGTGATCACTGATCTCCAAATTTTACAGATCAACAAAGGATCAAGAAAGGCTCAAG 770
DB 668 TTGTGATCACTGATCTCCAAATTTTACAGATCAACAAAGGATCAAGAAAGGCTCAAG 727
QY 771 TAAGCTGGGCTCTGCGAGAGGCAACGATCGCGGTGCGGAGGAGGCTTCGCGCTCT 830
DB 728 TAAGCTGGGCTCTGCGAGAGGCAACGATCGCGGTGCGGAGGAGGCTTCGCGCTCT 787
QY 831 TCGGCAACCTCTCTCTCTGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 890
DB 788 TCGGCAACCTCTCTCTCTGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 847
QY 891 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 950
DB 848 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 907
QY 951 TCTTCTGGGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1010
DB 908 TCTTCTGGGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 967
QY 1011 TGACATCTGTCAGAAATGAGTGAAGAAATTTTCTGCTCTCTCTCTCTCTCTCTCTCTCT 1070
DB 968 TGACATCTGTCAGAAATGAGTGAAGAAATTTTCTGCTCTCTCTCTCTCTCTCTCTCTCT 1027
QY 1071 GAGCCATTTTAAACAGACATCTGTCAGAAAGAAATGAGTGAAGTATTTCTGCTAA 1129
DB 1028 GAGCCATTTTAAACAGACATCTGTCAGAAAGAAATGAGTGAAGTATTTCTGCTAA 1086

```

```

RESULT 15
US-09-992-331-4
; Sequence 4, Application US/09992331
; Publication No. US20030022186A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, JOHN N.
; APPLICANT: MINTIER, GABE
; APPLICANT: RAMANATHAN, CHANDRA S.
; APPLICANT: HAWKEN, DONALD R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGRBWY18,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: D0048NP
; CURRENT APPLICATION NUMBER: US/09/992,331
; PRIOR APPLICATION NUMBER: 2001-11-14
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/261,782
; PRIOR FILING DATE: 2001-01-16

```

```

; PRIOR APPLICATION NUMBER: 60/248,483
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-992-331-4

```

```

Query Match 33.6%; Score 586; DB 10; Length 632;
Best Local Similarity 100.0%; Pred. No. 3.5e-277;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1130 TTTTCTTTTATACCGAGTTTCTACACCTGGGAGACTGTGGCATCTTTTAAACAGAT 1189
DB 1 TTTTCTTTTATACCGAGTTTCTACACCTGGGAGACTGTGGCATCTTTTAAACAGAT 60
QY 1190 TCATTTCAAGTACCTTCATCAGTGCACCCCTTTAAGAAATGAACCTATGCAATAG 1249
DB 61 TCATTTCAAGTACCTTCATCAGTGCACCCCTTTAAGAAATGAACCTATGCAATAG 120
QY 1250 ACATCCACAGCGTCGGTAAATTAAGGGTGAATCACCAGTTTCATATATTTTCCCTTTA 1309
DB 121 ACATCCACAGCGTCGGTAAATTAAGGGTGAATCACCAGTTTCATATATTTTCCCTTTA 180
QY 1310 TAAAGAGTTTGTGGCCAGGTGCAATGCTTCAATCCAGCAGTTTGGAGG 1369
DB 181 TAAAGAGTTTGTGGCCAGGTGCAATGCTTCAATCCAGCAGTTTGGAGG 240
QY 1370 CTGAGGTGGGTGATCAGCTGAGGTGAGAGTGAAGCAACCTGACCAATGAGT 1429
DB 241 CTGAGGTGGGTGATCAGCTGAGGTGAGAGTGAAGCAACCTGACCAATGAGT 300
QY 1430 ACCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1489
DB 301 ACCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 1490 AATCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1549
DB 361 AATCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 1550 GCAGTGAAGCGAGATGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1609
DB 421 GCAGTGAAGCGAGATGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 480
QY 1610 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1669
DB 481 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 540
QY 1670 TGTGTTGATATGATCAATTTAATAATTTTATTTATGACTGTT 1715
DB 541 TGTGTTGATATGATCAATTTAATAATTTTATTTATGACTGTT 586

```

```

Search completed: October 1, 2004, 04:22:19
Job time: 823.411 secs

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 19:03:02 ; Search time 5921.2 Seconds

(without alignments)
11419.163 Million cell updates/sec

Title: US-10-077-698-5

Perfect score: 1560

Sequence: 1 tctgccaagctcagcgtcgaagc.....aaaaaaaaaaaaaaaaa 1560

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenBank1: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_ov: *
7: gb_ph: *
8: gb_ph: *
9: gb_ph: *
10: gb_ph: *
11: gb_ph: *
12: gb_ph: *
13: gb_ph: *
14: gb_ph: *
15: gb_ph: *
16: gb_ph: *
17: gb_ph: *
18: gb_ph: *
19: gb_ph: *
20: gb_ph: *
21: gb_ph: *
22: gb_ph: *
23: gb_ph: *
24: gb_ph: *
25: gb_ph: *
26: gb_ph: *
27: gb_ph: *
28: gb_ph: *
29: gb_ph: *
30: gb_ph: *
31: gb_ph: *
32: gb_ph: *
33: gb_ph: *
34: gb_ph: *
35: gb_ph: *
36: gb_ph: *
37: gb_ph: *
38: gb_ph: *
39: gb_ph: *
40: gb_ph: *
41: gb_ph: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1560	100.0	1560	6	BD240721
2	1560	100.0	1560	6	BD269631
3	1560	100.0	1560	6	AR228217
4	1560	100.0	1560	6	AR372102
5	947	60.7	1419	10	BC053698
6	744	47.7	1086	10	AY288424
7	596	38.2	180944	2	AC112153
8	596	38.2	239718	2	AC101774
9	181	11.6	181	6	AR228218
10	138	8.8	138	6	AR228219
11	101	6.5	835	10	BC049237
12	92	5.9	169206	10	AC123550
13	92	5.9	277603	2	AC079543
14	77	4.9	212822	2	AC080156
15	77	4.9	225030	10	AC107608
16	77	4.9	247137	2	AC098268
17	77	4.9	249741	2	AC096330
18	77	4.9	271367	2	AC105467
19	77	4.9	271995	2	AC080155
20	57	3.7	238070	2	AC134313
21	57	3.7	242260	2	AC094567
22	57	3.7	255120	2	AC127219
23	51	3.3	187746	2	AC087123
24	49	3.1	271995	2	AC080155
25	44	2.8	744	9	AY255573
26	44	2.8	1086	6	AX498184
27	44	2.8	1321	6	AX168141
28	44	2.8	1458	6	BD015337
29	44	2.8	1737	6	AX549396
30	44	2.8	1743	6	BD240720
31	44	2.8	1743	6	BD269630
32	44	2.8	1743	6	AR228216
33	44	2.8	1743	6	AR372101
34	42	2.7	933	6	AX657504
35	40	2.6	1066	6	AX647273
36	40	2.6	1104	6	AX230123
37	40	2.6	1131	9	AY288417
38	40	2.6	163964	9	AL356214
39	39	2.5	638	6	AX277776
40	39	2.5	4291	9	HSX807867
41	38	2.4	203	9	AF43665882
42	38	2.4	241	6	BD210582
43	38	2.4	467	6	AX886416
44	38	2.4	487	6	BD026026
45	38	2.4	762	6	BD079639

ALIGNMENTS

RESULT 1
BD240721 1560 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION 14273 receptor, a G-protein coupled receptor.
ACCESSION
BD240721.1 GI:33050491
VERSION
BD240721.1 GI:33050491
KEYWORDS
JP 2002522011-A/2.
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1560)
Glucksmann, M.A. and Tsai, F.Y.
TITLE
14273 receptor, a G-protein coupled receptor
JOURNAL
Patent: JP 2002522011-A 2 23-JUL-2002;
MILLENNIUM PHARMACEUTICALS INC

COMMENT

OS Murine ortholog
 PN JP 2002522011-A/2
 PD 23-JUL-2002
 PR 30-JUN-1999 JP 2000557364
 PR 30-JUN-1998 US 09/107761, 30-DEC-1998 US 09/223538 PR
 26-FEB-1999 US 09/261599
 PI MARIA ALEXANDRA GLUCKSMANN, FONG YING TSAI
 PC C12N15/09, C07K14/705, C07K16/28, C12N1/15, C12N1/19, C12N1/21, PC
 C12N5/10,
 PC C12P21/02, C12P1/68, G01N33/15, G01N33/50, G01N33/53//C12P21/08,
 PC C12N15/00,
 CC C12N5/00
 CC 14273 receptor, a G-protein coupled receptor
 FH Key Location/Qualifiers
 FT source 1..1560
 FT Location/Qualifiers
 1..1560
 /organism="Murine ortholog",
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1560; DB 6; Length 1560;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCAAGCTCAGGTAAGCTCTTCCACTGCAATCTCAGAGAGGGGTTTCATGAGTGC 60
 DB 1 TTGGCAAGCTCAGGTAAGCTCTTCCACTGCAATCTCAGAGAGGGGTTTCATGAGTGC 60
 QY 61 TTACACCATCAGTACGACCTCCAGACTTGTCCGGCTTTTACCGGATCTTTCACAGCGGAG 120
 DB 61 TTACACCATCAGTACGACCTCCAGACTTGTCCGGCTTTTACCGGATCTTTCACAGCGGAG 120
 QY 121 TCGATGACCTCTTGGACAGCCAGCGCGCGAGCTCCGCACTCTTCCGCGAGCGTGG 180
 DB 121 TCGATGACCTCTTGGACAGCCAGCGCGCGAGCTCCGCACTCTTCCGCGAGCGTGG 180
 QY 121 TCGATGACCTCTTGGACAGCCAGCGCGCGAGCTCCGCACTCTTCCGCGAGCGTGG 180
 DB 121 TCGATGACCTCTTGGACAGCCAGCGCGCGAGCTCCGCACTCTTCCGCGAGCGTGG 180
 QY 181 GCGGGGCGCGGATGTCCTCTGATGTCGACAGAGAGAGCGGGCCCTGTCCTCGACA 240
 DB 181 GCGGGGCGCGGATGTCCTCTGATGTCGACAGAGAGAGCGGGCCCTGTCCTCGACA 240
 QY 181 GCGGGGCGCGGATGTCCTCTGATGTCGACAGAGAGAGCGGGCCCTGTCCTCGACA 240
 DB 181 GCGGGGCGCGGATGTCCTCTGATGTCGACAGAGAGAGCGGGCCCTGTCCTCGACA 240
 QY 241 CCTGAGCAAGTCAATCGACCCCACTTCCCTTCTCTGAGATGTCAAGGCGACACC 300
 DB 241 CCTGAGCAAGTCAATCGACCCCACTTCCCTTCTCTGAGATGTCAAGGCGACACC 300
 QY 301 GGTGGTGTGAGCGTGTGAGAGACACCGTTCTGGGACTCATCTTGTGCTCTACTGC 360
 DB 301 GGTGGTGTGAGCGTGTGAGAGACACCGTTCTGGGACTCATCTTGTGCTCTACTGC 360
 QY 361 TGGGCAACGTCGTCTCTAGTGTGTGGCGCCCTCGGCGCGCTGGGCGGTCAACCA 420
 DB 361 TGGGCAACGTCGTCTCTAGTGTGTGGCGCCCTCGGCGCGCTGGGCGGTCAACCA 420
 QY 421 GCGTGTGTCTCAACCTCTTCTGCGCGGATTTGCTCTTCAACAGGCGCATCCCTTAATGC 480
 DB 421 GCGTGTGTCTCAACCTCTTCTGCGCGGATTTGCTCTTCAACAGGCGCATCCCTTAATGC 480
 QY 481 TCGTCGTGCGCTGAGCGCTGAGCGCTGTTGGGCGCGTGTGCGACCTGCTCTTCT 540
 DB 481 TCGTCGTGCGCTGAGCGCTGAGCGCTGTTGGGCGCGTGTGCGACCTGCTCTTCT 540
 QY 541 ACGTGAATGACAAATAGGCGGAGCTGTCACATCTTCAACTGAGCGCGCTGAGAGC 600
 DB 541 ACGTGAATGACAAATAGGCGGAGCTGTCACATCTTCAACTGAGCGCGCTGAGAGC 600
 QY 601 GCATGTGTGATGTGTCGCTTCCGCGCGGCTTGAAGCGGCGCGGCGGAGCTCAAG 660
 DB 601 GCATGTGTGATGTGTCGCTTCCGCGCGGCTTGAAGCGGCGCGGCGGAGCTCAAG 660
 QY 661 CCGGCACTGTGCTTTTCAATAGGAGTTAATCGGCGCTGCGCGCTGCCCTTCAATCT 720
 DB 661 CCGGCACTGTGCTTTTCAATAGGAGTTAATCGGCGCTGCGCGCTGCCCTTCAATCT 720

DB 661 CCGGCACTGTGCTTTTCAATAGGAGTTAATCGGCGCTGCGCGCTGCCCTTCAATCT 720
 QY 721 TGTTCGCGGTGATCCGCGAGCGCTTCCGCGCGGAGACAGAAATTCGATTTGCACAT 780
 DB 721 TGTTCGCGGTGATCCGCGAGCGCTTCCGCGCGGAGACAGAAATTCGATTTGCACAT 780
 QY 781 TGAATGGCCCAACCGCATAGAGAAATCTCATGGATGTGTGTTTTTGAACCTTGAAC 840
 DB 781 TGAATGGCCCAACCGCATAGAGAAATCTCATGGATGTGTGTTTTTGAACCTTGAAC 840
 QY 841 TCTGTGTGCGGAGCTGTGTATGTGATGATGATCTTCAAAATTTTCAATGACGGAAG 900
 DB 841 TCTGTGTGCGGAGCTGTGTATGTGATGATGATCTTCAAAATTTTCAATGACGGAAG 900
 QY 901 CATGCGGAGAGAGCTTACGCTGAGCTGAGCATCTCTGAGAGCCACAGATCCGATGT 960
 DB 901 CATGCGGAGAGAGCTTACGCTGAGCTGAGCATCTCTGAGAGCCACAGATCCGATGT 960
 QY 961 CCCAACAAGACTACCGACTCTTCCGACGCTCTTCTGCTCATGTTTCTTTCATCA 1020
 DB 961 CCCAACAAGACTACCGACTCTTCCGACGCTCTTCTGCTCATGTTTCTTTCATCA 1020
 QY 1021 TGTGAGTCCCATCATCATCATCATCTCTCTCATCTTGAATCCAAACTTCCGAGAAC 1080
 DB 1021 TGTGAGTCCCATCATCATCATCATCTCTCTCATCTTGAATCCAAACTTCCGAGAAC 1080
 QY 1081 TGTGATCTGCGCATCTCTTCTTCTGAGGTGAGGCTTCAAGTTGCCACTGCCCC 1140
 DB 1081 TGTGATCTGCGCATCTCTTCTTCTGAGGTGAGGCTTCAAGTTGCCACTGCCCC 1140
 QY 1141 TAAACCCCATCTGTACACATGTGCTGTTACAGAAAGATGAGAGAAATTTTGTCT 1200
 DB 1141 TAAACCCCATCTGTACACATGTGCTGTTACAGAAAGATGAGAGAAATTTTGTCT 1200
 QY 1201 GCTTCTTTTTTCCAGAGAGAGAGCCATTTTACAGATAGCTCTGTAGCGGAAATGACT 1260
 DB 1201 GCTTCTTTTTTCCAGAGAGAGAGCCATTTTACAGATAGCTCTGTAGCGGAAATGACT 1260
 QY 1261 TGTCTGTTATTTCCAGCTAATCTAGGCTCTGTCGACAGTGAACACAGGCTGTCATGTA 1320
 DB 1261 TGTCTGTTATTTCCAGCTAATCTAGGCTCTGTCGACAGTGAACACAGGCTGTCATGTA 1320
 QY 1321 GGGAGTTAATCTTCAAGAGAAAGCCACCAAGTGCCTCTTAAATAACCCGACTTCA 1380
 DB 1321 GGGAGTTAATCTTCAAGAGAAAGCCACCAAGTGCCTCTTAAATAACCCGACTTCA 1380
 QY 1381 ACAGCAGGACTTCAAGAGAGAGCAATTAAGAAATATGCTCAGATATAAATAATTTT 1440
 DB 1381 ACAGCAGGACTTCAAGAGAGAGCAATTAAGAAATATGCTCAGATATAAATAATTTT 1440
 QY 1441 TCCCTTAAAGAACTTCTATAGGCTCTTCTTGTGAACCTTTTAAAGTGTGTAATAT 1500
 DB 1441 TCCCTTAAAGAACTTCTATAGGCTCTTCTTGTGAACCTTTTAAAGTGTGTAATAT 1500
 QY 1501 GATCTAGTTAATAATTTTATTTATTAACGCTTCTCAAAAAAATTTTTTTTTT 1560
 DB 1501 GATCTAGTTAATAATTTTATTTATTTATTAACGCTTCTCAAAAAAATTTTTTTTTT 1560

RESULT 2
 BD269631
 LOCUS 1560 bp DNA linear PAT 17-JUL-2003
 DEFINITION 14273 receptor, a novel G-protein coupled receptor.
 ACCESSION BD269631.1 GI:33079399
 VERSION BD269631.1
 KEYWORDS JP 2002536997-A/2.
 SOURCE
 Mus sp.
 ORGANISM Mus sp.
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1560)
 Glucksmann, M.A. and Tsai, F.Y.
 14273 receptor, a novel G-protein coupled receptor

JOURNAL Patent: US 6448005-A 5 10-SEP-2002;
 FEATURES location/Qualifiers
 source 1..1560
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1560; DB 6; Length 1560;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTGCGAAGCTCAGCGTAAGCCTCTTCCACTGCAATCTCAGAGAAAGGGGTTTCATGAGTGC 60
DB 1 TTGCGAAGCTCAGCGTAAGCCTCTTCCACTGCAATCTCAGAGAAAGGGGTTTCATGAGTGC 60
QY 61 TTACACCAATCAGAGACCACTCCAGACTTGTCCGGCTTTACCCGAATCTTCAGACGGAG 120
DB 61 TTACACCAATCAGAGACCACTCCAGACTTGTCCGGCTTTACCCGAATCTTCAGACGGAG 120
QY 121 TCGATGACCTCTTGAAGCCAGAGCGCGAGCTCCGCACTCTCCGGAAGCGGTGG 180
DB 121 TCGATGACCTCTTGAAGCCAGAGCGCGAGCTCCGCACTCTCCGGAAGCGGTGG 180
QY 181 GCCGGGCGCCGCGATGCTCCCTGAGTGTGCAAGACGAGCGGCCCTGATCTCCGACA 240
DB 181 GCCGGGCGCCGCGATGCTCCCTGAGTGTGCAAGACGAGCGGCCCTGATCTCCGACA 240
QY 241 CCTCGACCAAGTCATTCGACCACTTCCTCTTCTTCTGAGATGTCAAGGGCCACACC 300
DB 241 CCTCGACCAAGTCATTCGACCACTTCCTCTTCTTCTGAGATGTCAAGGGCCACACC 300
QY 301 GGTGAGGTGAGAGCGTGTGAGACACCGTTCTGAGGACTCATCTTGTGTGCTCACTGC 360
DB 301 GGTGAGGTGAGAGCGTGTGAGACACCGTTCTGAGGACTCATCTTGTGTGCTCACTGC 360
QY 361 TGGGCAACGTGTGTCTCTAGTGTGTGCGCGCGCTGCGGCGCGTGGGGCGTCAAGCA 420
DB 361 TGGGCAACGTGTGTCTCTAGTGTGTGCGCGCGCTGCGGCGCGTGGGGCGTCAAGCA 420
QY 421 GCTGTGTCTCAACCTTTCTTGTGCGGAGTTTGTCTTTCACAGGCGCATCTCTTATGTG 480
DB 421 GCTGTGTCTCAACCTTTCTTGTGCGGAGTTTGTCTTTCACAGGCGCATCTCTTATGTG 480
QY 481 TCGTGTGCTGTGAGTGTGAGGCTGTGCTGTGAGGAGCGCTGTGTGCAACCTGTCTTCT 540
DB 481 TCGTGTGCTGTGAGTGTGAGGCTGTGCTGTGAGGAGCGCTGTGTGCAACCTGTCTTCT 540
QY 541 ACGTGATGACATGAGCGGAGCGTCAAGATCTTCACACTTGGCGGCTGAGCTGTGAGC 600
DB 541 ACGTGATGACATGAGCGGAGCGTCAAGATCTTCACACTTGGCGGCTGAGCTGTGAGC 600
QY 601 GCATGTGTGATGTGTGCTGCTGCGGCGCGCTTGAACGCGCGCGGCGCGGAGCTGAG 660
DB 601 GCATGTGTGATGTGTGCTGCTGCGGCGCGCTTGAACGCGCGCGGCGCGGAGCTGAG 660
QY 661 CGGACACTGTGCTTTCATATGAGGTTACTGCGGCTCTGCGGCTGCGGCTCTTCAATCT 720
DB 661 CGGACACTGTGCTTTCATATGAGGTTACTGCGGCTCTGCGGCTGCGGCTCTTCAATCT 720
QY 721 TGTTCGCGTGTGCTGCGGAGCGCTTCCGCGGCGGAGCAAGAAATTCGATTTGCAAT 780
DB 721 TGTTCGCGTGTGCTGCGGAGCGCTTCCGCGGCGGAGCAAGAAATTCGATTTGCAAT 780
QY 781 TGAATTTGGCCCAACCGGATGAGAAATTCATATGGATGTGTTTTTGAAGCTTTGAACT 840
DB 781 TGAATTTGGCCCAACCGGATGAGAAATTCATATGGATGTGTTTTTGAAGCTTTGAACT 840
QY 841 TCCGTGTGCGGAGCTGTGATCTTGTGATGAGTTACTCAAAATTTTACAGATCGAAG 900
DB 841 TCCGTGTGCGGAGCTGTGATCTTGTGATGAGTTACTCAAAATTTTACAGATCGAAG 900
QY 901 CATGCGGAAGAGCTTACGCTGAGCTTGGCATCTGAGAGCCACAGATCCGAGTGT 960
DB 901 CATGCGGAAGAGCTTACGCTGAGCTTGGCATCTGAGAGCCACAGATCCGAGTGT 960

```

```

DB 901 CATGCGGAAGAGCTTACGCTGAGCTTGGCATCTCTGAGAGCCACAGATCCGAGTGT 960
QY 961 CCCAACAAGACTACCGACTCTTCCGAGAGCTCTTCCGACTCATGATGTTCTTTCATCA 1020
DB 961 CCCAACAAGACTACCGACTCTTCCGAGAGCTCTTCCGACTCATGATGTTCTTTCATCA 1020
QY 1021 TGTGAGTCCCATCATCATCATCATCATCTCTCTCATCTTGTATCCAAATCTCCGAGAGC 1080
DB 1021 TGTGAGTCCCATCATCATCATCATCATCTCTCTCATCTTGTATCCAAATCTCCGAGAGC 1080
QY 1081 TGTGATCTGCGCATCTCTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1140
DB 1081 TGTGATCTGCGCATCTCTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1140
QY 1141 TAAACCCATCTGTAACCAATGCTGCTTTCAGAAAGCAATGAGAGAAATTTTGTCT 1200
DB 1141 TAAACCCATCTGTAACCAATGCTGCTTTCAGAAAGCAATGAGAGAAATTTTGTCT 1200
QY 1201 GCTTCTTTTTCAGAGAAAGGAGCCATTTTTCAGATACGCTGTGAGGCGAAATGACT 1260
DB 1201 GCTTCTTTTTCAGAGAAAGGAGCCATTTTTCAGATACGCTGTGAGGCGAAATGACT 1260
QY 1261 TGTCTGTTATTTTCAGACTTACGCTCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1320
DB 1261 TGTCTGTTATTTTCAGACTTACGCTCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1320
QY 1321 GGGAGTTAATCTTCAAGAAAGCCACAGATGCGCTGCTTAAATAATCCGACTTCA 1380
DB 1321 GGGAGTTAATCTTCAAGAAAGCCACAGATGCGCTGCTTAAATAATCCGACTTCA 1380
QY 1381 ACAGAGGATCTACGAGAGCCAGCAATTAAGAAATGATGCTCAGTATATAATTTT 1440
DB 1381 ACAGAGGATCTACGAGAGCCAGCAATTAAGAAATGATGCTCAGTATATAATTTT 1440
QY 1441 TCTTAAATAAATCTTCTATGAGTGTCTTGTGAACTTTTAAAGTGTGTAATAT 1500
DB 1441 TCTTAAATAAATCTTCTATGAGTGTCTTGTGAACTTTTAAAGTGTGTAATAT 1500
QY 1501 GATCTAGTTAATAATTTTAAATTTTAACTGTTCTTCAAAAAAATTTTTTTTTT 1560
DB 1501 GATCTAGTTAATAATTTTAAATTTTAACTGTTCTTCAAAAAAATTTTTTTTTT 1560

```

RESULT 4
 AR372102 1560 bp DNA linear PAT 12-SEP-2003
 LOCUS AR372102
 DEFINITION Sequence 5 from patent US 6395877.
 ACCESSION AR372102
 VERSION AR372102.1 GI:34609379
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1560)
 AUTHORS Glucksmann, M.A. and Tsai, F.-Y.
 TITLE 14273 receptor, a novel G-protein coupled receptor
 JOURNAL Patent: US 6395877-A 5 28-MAY-2002;
 FEATURES location/Qualifiers
 source 1..1560
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1560; DB 6; Length 1560;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTGCGAAGCTCAGCGTAAGCCTCTTCCACTGCAATCTCAGAGAAAGGGGTTTCATGAGTGC 60
DB 1 TTGCGAAGCTCAGCGTAAGCCTCTTCCACTGCAATCTCAGAGAAAGGGGTTTCATGAGTGC 60
QY 61 TTACACCAATCAGAGACCACTCCAGACTTGTCCGGCTTTACCCGAATCTTCAGACGGAG 120
DB 61 TTACACCAATCAGAGACCACTCCAGACTTGTCCGGCTTTACCCGAATCTTCAGACGGAG 120

```

Db 61 TTCACACATCATGATGACCACTCAGACCTTGTCCGGCTTACCCGAATCTTCACAGCGAG 120
 Qy 121 TCGATGACCTCTTTGACAGCCAGAGCGCGGCGAGCTCCGCATCTTCCGGAGCGGTGG 180
 Db 121 TCGATGACCTCTTTGACAGCCAGAGCGCGGCGAGCTCCGCATCTTCCGGAGCGGTGG 180
 Qy 181 GCGGAGCGCGGAGATGCCCCGAGTGTGACAGAGAGAGGCGCGCTGTGTCCCTCGGACA 240
 Db 181 GCGGAGCGCGGAGATGCCCCGAGTGTGACAGAGAGAGGCGCGCTGTGTCCCTCGGACA 240
 Qy 241 CCCTGACCAAGTCAATCGACCCACTTCTCTCGAGATGTCAGAGGCGAGCACC 300
 Db 241 CCCTGACCAAGTCAATCGACCCACTTCTCTCGAGATGTCAGAGGCGAGCACC 300
 Qy 301 GGTGTGTGTGAGCGGTGTGAGACCAACCGTTCTGGAGCTCATCTTTGTCTTCACTGC 360
 Db 301 GGTGTGTGTGAGCGGTGTGAGACCAACCGTTCTGGAGCTCATCTTTGTCTTCACTGC 360
 Qy 361 TGGGCAACGATGTGCTCTAGTGTGAGGCGCGCGCTGGCGCGCGGTGGCGCGTCA 420
 Db 361 TGGGCAACGATGTGCTCTAGTGTGAGGCGCGCGCTGGCGCGCGGTGGCGCGTCA 420
 Qy 421 GCGTGTGTCAACCTTCTGCGCGGATTTGCTTTCACAGAGCGCATCCCTTAGTGC 480
 Db 421 GCGTGTGTCAACCTTCTGCGCGGATTTGCTTTCACAGAGCGCATCCCTTAGTGC 480
 Qy 481 TCGTGTGCGCTGGACTGAGGCGCTGCTGTGGGGCCGTCGTGCACTGCTTCT 540
 Db 481 TCGTGTGCGCTGGACTGAGGCGCTGCTGTGGGGCCGTCGTGCACTGCTTCT 540
 Qy 541 ACCTGATGACAAATGAGGCGGAGCGCTACAGATCTTCAACATGCGCGCGTCACTGAGC 600
 Db 541 ACCTGATGACAAATGAGGCGGAGCGCTACAGATCTTCAACATGCGCGCGTCACTGAGC 600
 Qy 601 GCATGTGTGATCATGTGCGCTCCGCGCGGCTTTGAGCGCGCGCGCGCGGAGCTGAG 660
 Db 601 GCATGTGTGATCATGTGCGCTCCGCGCGGCTTTGAGCGCGCGCGCGCGGAGCTGAG 660
 Qy 661 CCGCACTGTGCTTCAATGAGGGTACTGCGCGCTGCGCGCGCTGCTTCACT 720
 Db 661 CCGCACTGTGCTTCAATGAGGGTACTGCGCGCTGCGCGCGCTGCTTCACT 720
 Qy 721 TGTTCGCGGTGTGCGCGAGCGCTTCCGCGCGGAGCAAGAAATTCGATTTGACAT 780
 Db 721 TGTTCGCGGTGTGCGCGAGCGCTTCCGCGCGGAGCAAGAAATTCGATTTGACAT 780
 Qy 781 TGTGATGCGCCCAATGAGAGAAATCATGAGATGTGTTTGTGAGCTTGAAT 840
 Db 781 TGTGATGCGCCCAATGAGAGAAATCATGAGATGTGTTTGTGAGCTTGAAT 840
 Qy 841 TCTGTGTGCGGAGCTGTGATGTGATCAATCAATTTTACAGATCAGAAAG 900
 Db 841 TCTGTGTGCGGAGCTGTGATGTGATCAATCAATTTTACAGATCAGAAAG 900
 Qy 901 CATCGCGGAAAGAGCTTACCTGAGCTGTGAGTACTTGAAGCAGCAAGTCCGAGTGT 960
 Db 901 CATCGCGGAAAGAGCTTACCTGAGCTGTGAGTACTTGAAGCAGCAAGTCCGAGTGT 960
 Qy 961 CCCAACAAGACTACCGACTCTTCGCGAGCGCTTCCGCTCATGAGTTCCTTCACTCA 1020
 Db 961 CCCAACAAGACTACCGACTCTTCGCGAGCGCTTCCGCTCATGAGTTCCTTCACTCA 1020
 Qy 1021 TGTGAGTCCCATCATCATACCAATCTCTCATCTTGTATCAAAATTCGCGAGAGC 1080
 Db 1021 TGTGAGTCCCATCATCATACCAATCTCTCATCTTGTATCAAAATTCGCGAGAGC 1080
 Qy 1081 TGTGTATCTGAGCATCCCTTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
 Db 1081 TGTGTATCTGAGCATCCCTTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
 Qy 1141 TAAACCCCATCTGTATCAACATGTCTGTTTCAGAGAAAGAAATTTTGTCT 1200
 Db 1141 TAAACCCCATCTGTATCAACATGTCTGTTTCAGAGAAAGAAATTTTGTCT 1200

Qy 1201 GCTTCTTTTTCAGAGAGGAGCCATTTTATAGATAGCTCTGTGAGGGAATGACT 1260
 Db 1201 GCTTCTTTTTCAGAGAGGAGCCATTTTATAGATAGCTCTGTGAGGGAATGACT 1260
 Qy 1261 TGTCTGTTATTTCCAGCTTAATAGCTCTGTGTGAGAGAGGAGGAGGAGGAGGAGGAG 1320
 Db 1261 TGTCTGTTATTTCCAGCTTAATAGCTCTGTGTGAGAGAGGAGGAGGAGGAGGAGGAG 1320
 Qy 1321 GGGAGTTAACTTCAAGAGAAAGCCAGCAGTGCCTGCTTTTAAATAACCCGACTTCA 1380
 Db 1321 GGGAGTTAACTTCAAGAGAAAGCCAGCAGTGCCTGCTTTTAAATAACCCGACTTCA 1380
 Qy 1381 ACAGCAGGATCTACGAGAGCCAGAAATTAAGAAATGATGCTCATATATAATTTT 1440
 Db 1381 ACAGCAGGATCTACGAGAGCCAGAAATTAAGAAATGATGCTCATATATAATTTT 1440
 Qy 1441 TCTTAAAGAACTTTATAGGTTCTTTTGTGAACTTTTATAGTGTGTATPAT 1500
 Db 1441 TCTTAAAGAACTTTATAGGTTCTTTTGTGAACTTTTATAGTGTGTATPAT 1500
 Qy 1501 GATCTAGTTAATAATTTTATATATATATACGTGTTCTTCAAAAAAATTTTAAAA 1560
 Db 1501 GATCTAGTTAATAATTTTATATATATATATACGTGTTCTTCAAAAAAATTTTAAAA 1560
 RESULT 5
 BC053698
 LOCUS
 DEFINITION Mus musculus G protein-coupled receptor 120, mRNA (CDNA clone
 ACCSSION BC053698
 VERSION BC053698.1 GI:31566089
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
 1 (bases 1 to 1419)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hale, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 JOURNAL MEDLINE
 PUBMED 12477932
 2 (bases 1 to 1419)
 Strausberg, R.
 Reference
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hghri.nih.gov
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stancilpop,S., Thomas,P.J., Touchman,J.W.,
Tsouroun,C., Vogt,J.L., Walker,M.A., Wecherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>
Series: IRAC Plate: 116 Row: n Column: 18.

FEATURES
source

1. 1419
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:62870 IMAGE:6529120"
/issue_type="Colon, normal, 5 month old male mouse."
/clone_lib="NCI CGAP_Co24"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1419
/gene="Gpr120"
/db_xref="LocusID:107221"
/db_xref="MGI:2147577"
49. 1134
/codon_start=1
/product="G protein-coupled receptor 120"
/protein_id="AAH53698.1"
/db_xref="GI:31566090"
/db_xref="LocusID:107221"
/translation="MSPRCQOTTEPGSHLDVQNRTHFPSPSVKXKDHVLVSIVET
TVGLLFVSLGNVCLVIVARRRRGATVSLVNFCDLFTSAIPLVIVRWTE
AMIGPVCVHLFPVMTSGSVYILTLAAVSLRMVCIVARRELSPGRRTAALIA
FIVGYSALALPLCILFRVVPORLPGSDOEIPICITLMDPRLIGISINDVFEVTLNFIY
PGIVIVSYKILIOITKASRKRLTLISAYESHQIRASOODYRLFRLLPLFMNSFFIM
WSPILITLILIONFERODLVIMPSLFPFWVAITFANSAINPLIYNMSLFRNMRKIF
CCFPPEKGAIFIDTISVRNDLSVITS"
277. 1011
/note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
family)"
/db_xref="CDD:pfam00001"

gene
CDS

misc_feature
277. 1011
/note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
family)"
/db_xref="CDD:pfam00001"

ORIGIN

Query Match 60.7%; Score 947; DB 10; Length 1419;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1247; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 231 CCTCGGACACCCCGGACCAAGTACGACCCGACTTCCCTTCTCTCGGAGTGAAG 290
DB 85 CCTCGGACACCCCGGACCAAGTACGACCCGACTTCCCTTCTCTCGGAGTGAAG 144
QY 291 GGGACACACCGGTGTGTGAGCGTGTGAGAGACCAACCGTCTTGAGACTGATTTTGTG 350
DB 145 GGGACACACCGGTGTGTGAGCGTGTGAGAGACCAACCGTCTTGAGACTGATTTTGTG 204
QY 351 GTCTCACTGTGTGGGCAAGTGTGTCTTCTAGTGTGTGTGGCGCGCGCGTGGCGCGTGGG 410
DB 205 GTCTCACTGTGTGGGCAAGTGTGTCTTCTAGTGTGTGTGGCGCGCGCGTGGCGCGTGGG 264
QY 411 GCGTCAACGACCGTGTGTCTTCTTCTGCGCGGATTTGCTTTCACCAAGGCGCATC 470
DB 265 GCGTCAACGACCGTGTGTCTTCTTCTGCGCGGATTTGCTTTCACCAAGGCGCATC 324

QY 471 CTTCTAGTGTCTGTCTGTGCGCTGAGTCTGAGGCGCTGTGTGGGCGCCGCTGTGCGAC 530
DB 325 CTTCTAGTGTCTGTCTGTGCGCTGAGTCTGAGGCGCTGTGTGGGCGCCGCTGTGCGAC 384
QY 531 CTGCTCTTCAAGTGAACAAATGAGCGGCGTCAAGTCTCTCACTGAGCGCGGTC 590
DB 385 CTGCTCTTCAAGTGAACAAATGAGCGGCGTCAAGTCTCTCACTGAGCGCGGTC 444
QY 591 AGCTGAGGCGCATGTGTGTCATGTGCGCTTCCGCGCGGCTTGAAGCGGCGCGGCGG 650
DB 445 AGCTGAGGCGCATGTGTGTCATGTGCGCTTCCGCGCGGCTTGAAGCGGCGCGGCGG 504
QY 651 CGGACTGAGCGGCACTGTGCTGCTTTCATATGGGGTTACTGGCGCTGCGCGGCTGCC 710
DB 505 CGGACTGAGCGGCACTGTGCTGCTTTCATATGGGGTTACTGGCGCTGCGCGGCTGCC 564
QY 711 CTCTACATCTGTGTCCGCGGTGTCGCGGCGGCGCTTCCGCGCGGCGGCGGCGGCGG 770
DB 565 CTCTGATCTGTGTTCGCGGTGTCGCGGCGGCGCTTCCGCGCGGCGGCGGCGGCGG 624
QY 771 ATTGACATTTGATTGGCCCAACCGCATAGAGAAATCTCATGGAGTGTGTTTTAG 830
DB 625 ATTGACATTTGATTGGCCCAACCGCATAGAGAAATCTCATGGAGTGTGTTTTAGT 684
QY 831 ACTTTGAATCTTCTGTCGCGGCGGCGATGTGTGATCATGTTACTTCCAAATTTTACG 890
DB 685 ACTTTGAATCTTCTGTCGCGGCGGCGGCGATGTGTGATCATGTTACTTCCAAATTTTACG 744
QY 891 ATACGAAACATCGGCGGAGAGGCTTAGCGTAGACTGTGACATCTACTGTAGAGCACAG 950
DB 745 ATACGAAACATCGGCGGAGAGGCTTAGCGTAGACTGTGACATCTACTGTAGAGCACAG 804
QY 951 ATCCGAGTGTCCCAACAAGACTACCGACTCTTCGCGCGCTCTCTCATGTGTTTC 1010
DB 805 ATCCGAGTGTCCCAACAAGACTACCGACTCTTCGCGCGCTCTCTCATGTGTTTC 864
QY 1011 TTCTTCATCATGTGAGTCCCATCATCATCATCATCTCTCTCATGTGTTTC 1070
DB 865 TTCTTCATCATGTGAGTCCCATCATCATCATCATCTCTCTCATGTGTTTC 924
QY 1071 CGGACAGACCTGTGATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1130
DB 925 CGGACAGACCTGTGATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 984
QY 1131 AACTGCGCTTAAACCCCATCTGTAACAATGTGTGTGTGAGGAGCGAATGAGAGAG 1190
DB 985 AACTGCGCTTAAACCCCATCTGTAACAATGTGTGTGTGAGGAGCGAATGAGAGAG 1044
QY 1191 ATTTTGTGTGCTTCTTTTTCAGAGAGAGGAGCCATTTTACATAGTGTGTGAGG 1250
DB 1045 ATTTTGTGTGCTTCTTTTTCAGAGAGAGGAGCCATTTTACATAGTGTGTGAGG 1104
QY 1251 CGAAATGACTGTGTGTATTTTTCAGGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1310
DB 1105 CGAAATGACTGTGTGTATTTTTCAGGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1164
QY 1311 TGCAATGAAAGGAGTTAATCTTCAAGGAAAGCCACCAAGTGTGTGTGTGTGTGTGTGT 1370
DB 1165 TGCAATGAAAGGAGTTAATCTTCAAGGAAAGCCACCAAGTGTGTGTGTGTGTGTGTGT 1224
QY 1371 CGGACTTCAACAGAGGCACTTACGAGGCGGAGCAATTAAGAAATATGTGTGTGTGTGT 1430
DB 1225 CGGACTTCAACAGAGGCACTTACGAGGCGGAGCAATTAAGAAATATGTGTGTGTGTGT 1284
QY 1431 AAAATATTTTCTTAAAGAACTTCTAAGGGTCCCTTTGTGAACTTTT 1483
DB 1285 AAAATATTTTCTTAAAGAACTTCTAAGGGTCCCTTTGTGAACTTTT 1337

RESULT 6
AY288424 1086 bp mRNA linear ROD 08-DEC-2003
LOCUS AY288424
DEFINITION Mus musculus G protein-coupled receptor 120 (Gpr120) mRNA, complete

[illegible]

OY		591	AGCTGGAGCCCATGATGTGCATGTGGCGCTTCGGCGCGGCGCTTGAGCGGCCCGGGCGG	650
Db		337	AGCTGGAGCCCATGTGTGCATGTGGCGCTTCGGCGCGGCGCTTGAGCGGCCCGGGCGG	456
OY		651	CGGACTCAGGCGGCACTGTGCTGCTTTCATATAGGGTTACTCGGCGCTTCGGCGCGTCC	710
Db		457	CGGACTCAGGCGGCACTGTGCTGCTTTCATATAGGGTTACTCGGCGCTTCGGCGCGTCC	516
OY		711	CTTACATCTTGTTCCGGTGTGTCCCGAGGCGCTTCCC GGCGGGGACCCAGAAATTCCG	770
Db		517	CTGTGCATCTTGTTCCGGTGTGTCCCGAGGCGCTTCCC GGCGGGGACCCAGAAATTCCG	576
OY		771	ATTGACATTTGGATTTGGCCCCAACCGCATAGAGAATAATCATGGAGATGTTTTTTGAG	830
Db		577	ATTGACATTTGGATTTGGCCCCAACCGCATAGAGAATAATCATGGAGATGTTTTTTGAG	636
OY		831	ACTTTGAACCTTCGTGTGCGCGGACCTGTCTATTGTGATCAAGTTACTCCAAATTTTACG	890
Db		637	ACTTTGAACCTTCGTGTGCGCGGACCTGTCTATTGTGATCAAGTTACTCCAAATTTTACG	696
OY		891	ATCACGAAGAATGCGCGGAGAGGCTTAACGCTGAGCTTGGCATACTCTGAGAGCCACCAG	950
Db		697	ATCACGAAGAATGCGCGGAGAGGCTTAACGCTGAGCTTGGCATACTCTGAGAGCCACCAG	756
OY		951	ATCCGAGTGTCCCAAACAAGACTACCGACTCTCCGACGCGCTCTCCGTCAATGGTTCC	1010
Db		757	ATCCGAGTGTCCCAAACAAGACTACCGACTCTCCGACGCGCTCTCCGTCAATGGTTCC	816
OY		1011	TTCCTCATCATGTGGAGTCCCATCATCATCAATCCCTCATCTTGATCCAAAATTCC	1070
Db		817	TTCCTCATCATGTGGAGTCCCATCATCATCAATCCCTCATCTTGATCCAAAATTCC	876
OY		1071	CGGAGAGACTGTGATCTGGCCATCCCTTTCTTCTGTGGTGTGGCTTCAGCTTTGCC	1130
Db		877	CGGAGAGACTGTGATCTGGCCATCCCTTTCTTCTGTGGTGTGGCTTCAGCTTTGCC	936
OY		1131	AACCTGTGCTTAAACCCCATACTGTACAACTATGTGCTGTTCAGAGAACGATGAGAGAG	1190
Db		937	AACCTGTGCTTAAACCCCATACTGTACAACTATGTGCTGTTCAGAGAACGATGAGAGAG	996
OY		1191	ATTTTGTGCTGCTCTTTTTCGAGAGAGGAGCCATTTTACAGATACGCTGTGCAGS	1250
Db		997	ATTTTGTGCTGCTCTTTTTCGAGAGAGGAGCCATTTTACAGATACGCTGTGTGCAGS	1056
OY		1251	GGAATGACTTGTCTGTATTTCAGCTAA	1280
Db		1057	GGAATGACTTGTCTGTATTTCAGCTAA	1086
RESULT 7				
AC112153/c				
LOCUS		180944 bp	DNA	linear ROD 27-NOV-2003
DEFINITION		Mus musculus BAC clone RP24-93G11 from chromosome 19, complete		
ACCESSION		AC112153		
VERSION		AC112153.4	GI:34495034	
KEYWORDS		HTG.		
SOURCE		Mus musculus (house mouse)		
ORGANISM		Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE		Tomlinson,C. and Haglund,K.		
JOURNAL		The sequence of Mus musculus BAC clone RP24-93G11		
REFERENCE		Unpublished (2001)		
AUTHORS		Wilson,R.		
TITLE		Sequencing of Mus musculus		
JOURNAL		Unpublished (2001)		
REFERENCE		3 (bases 1 to 180944)		
AUTHORS		McPherson,J.D. and Waterston,R.H.		
TITLE		Direct Submission		
JOURNAL		Submitted (19-FEB-2002) Genome Sequencing Center, 4444 Forest Park		

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (15-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 180944)
Wilson, R.K.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (06-SEP-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 180944)
Wilson, R.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (27-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 6, 2003 this sequence version replaced g130725964.

Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu

Summary Statistics
Center project name: M_BB0093G11

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1..180944
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/map="19"
/clone="RP24-93G11"
/clone_1id="RPCI-24"
266..2347
/rpt_family="B4"
repeat_region
3134..3197
/rpt_family="MIR"
repeat_region
4136..4276
/rpt_family="MIR"
repeat_region
6494..6605
/rpt_family="MIR"
repeat_region
7102..7240
/rpt_family="Alu"
repeat_region
7468..7643

repeat_region
7871..8016
/rpt_family="Alu"
repeat_region
8056..8245
/rpt_family="B2"
repeat_region
8673..8846
/rpt_family="MIR"
repeat_region
10233..10292
/rpt_family="L1"
repeat_region
10313..10440
/rpt_family="Alu"
repeat_region
11684..11817
/rpt_family="MIR"
repeat_region
12708..13016
/rpt_family="ERVK"
repeat_region
13479..13623
/rpt_family="B4"
repeat_region
14203..14269
/rpt_family="MALR"
repeat_region
15271..15372
/rpt_family="Alu"
repeat_region
16088..16476
/rpt_family="MALR"
repeat_region
16482..16595
/rpt_family="MIR"
unsure
17326..17397
/note="Sequence derived from one plasmid subclone."
18297..18499
/rpt_family="ERVK"
22015..22065
/rpt_family="MER1_type"
22082..22236
/rpt_family="ERVK"
22399..22457
/rpt_family="MER1_type"
22461..22756
/rpt_family="B4"
22783..22894
/rpt_family="MER1_type"
23296..23566
/rpt_family="MALR"
23567..23713
/rpt_family="Alu"
23714..23906
/rpt_family="MALR"
24494..24544
/rpt_family="MIR"
24545..24685
/rpt_family="B4"
24700..24863
/rpt_family="MIR"
25002..25052
/rpt_family="ERV1"
25062..25155
/rpt_family="MER2_type"
25212..25626
/rpt_family="RMR10A"
25797..26037
/rpt_family="MER2_type"
26074..26498
/rpt_family="RMR10A"
26526..26657
/rpt_family="MER2_type"
26706..26818
/rpt_family="MER1_type"
28012..28300
/rpt_family="ERV1"
28302..28543
/rpt_family="L1"
29192..29283
/rpt_family="ERVK"
29861..29902
/rpt_family="L1"

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: 113J_17
 Center clone name: 113J_17

----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 23359 bases at least Q40
 Consensus quality: 236645 bases at least Q30
 Consensus quality: 236645 bases at least Q20
 Insert size: 21000; agarose-fp
 Insert size: 237818; sum-of-coverage
 Quality coverage: 9.1 in Q20 bases; agarose-fp
 Quality coverage: 8.1 in Q20 bases; sum-of-coverage

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 20 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      309: contig of 309 bp in length
*      310      409: gap of 100 bp
*      410      1476: contig of 1067 bp in length
*      1477      1576: gap of 100 bp
*      1577      2753: contig of 1177 bp in length
*      2754      2853: gap of 100 bp
*      2854      3880: contig of 1027 bp in length
*      3881      3980: gap of 100 bp
*      3981      5036: contig of 1056 bp in length
*      5037      5135: gap of 100 bp
*      5137      6296: contig of 1160 bp in length
*      6297      6396: gap of 100 bp
*      6397      7884: contig of 1488 bp in length
*      7885      7984: gap of 100 bp
*      7985      9674: contig of 1690 bp in length
*      9675      9774: gap of 100 bp
*      9775      10981: contig of 1207 bp in length
*      10982      11081: gap of 100 bp
*      11082      13247: contig of 2166 bp in length
*      13248      13347: gap of 100 bp
*      13348      16734: contig of 3387 bp in length
*      16735      16834: gap of 100 bp
*      16835      22639: contig of 6705 bp in length
*      22640      23540: gap of 100 bp
*      23540      29386: contig of 5747 bp in length
*      29387      29486: gap of 100 bp
*      29487      43815: contig of 14329 bp in length
*      43816      43915: gap of 100 bp
*      43916      55354: contig of 11439 bp in length
*      55355      55454: gap of 100 bp
*      55455      82497: contig of 27043 bp in length
*      82498      82597: gap of 100 bp
*      82598      109798: contig of 27201 bp in length
*      109799      109898: gap of 100 bp
*      109899      181806: contig of 71908 bp in length
*      181807      213173: contig of 31267 bp in length
*      213174      213273: gap of 100 bp
*      213274      239718: contig of 26445 bp in length.
  
```

Location/Qualifiers
 1. .239718
 /organism="Mus musculus"
 /mol_type="genomic DNA"

```

/misc_feature      /db_xref="taxon:10090"
                   /clone="RP24-113J17"
                   /clone_1b="RP24 Male Mouse BAC"
1. .309
  /note="assembly_fragment"
  /clone_end:Spe
  vector_side:left"
  410. .1476
  /note="assembly_fragment"
  /note="assembly_fragment"
  1577. .2753
  /note="assembly_fragment"
  2854. .3880
  /note="assembly_fragment"
  3981. .5036
  /note="assembly_fragment"
  5137. .6296
  /note="assembly_fragment"
  6397. .7884
  /note="assembly_fragment"
  7985. .9674
  /note="assembly_fragment"
  9775. .10981
  /note="assembly_fragment"
  11082. .13247
  /note="assembly_fragment"
  13348. .16734
  /note="assembly_fragment"
  16835. .23539
  /note="assembly_fragment"
  23640. .29386
  /note="assembly_fragment"
  29487. .43815
  /note="assembly_fragment"
  43916. .55354
  /note="assembly_fragment"
  55455. .82497
  /note="assembly_fragment"
  82598. .109798
  /note="assembly_fragment"
  109899. .181806
  /note="assembly_fragment"
  181907. .213173
  /note="assembly_fragment"
  213274. .239718
  /note="assembly_fragment"
  clone_end:T7
  vector_side:right"
  
```

ORIGIN

Query Match 38.2%; Score 596; DB 2; Length 239718;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      888 CAGATCCGGAAGATCGCGAAGAGCTTACGCTTGACATCTGTGAGGCAC 947
         |||||
Db      100558 CAGATCCGGAAGATCGCGAAGAGCTTACGCTTGACATCTGTGAGGCAC 100617

QY      948 CAGATCCGAGTGTCCCAAGACTACGACTTCCGACGCTCTTCTGATGATGTT 1007
         |||||
Db      100618 CAGATCCGAGTGTCCCAAGACTACGACTTCCGACGCTCTTCTGATGATGTT 100677

QY      1008 TCCTTCTTCATGTGAGTCCCATCATCATCCATCTCTGATCTTGAATCAAAAC 1067
         |||||
Db      100678 TCCTTCTTCATGTGAGTCCCATCATCATCCATCTCTGATCTTGAATCAAAAC 100737

QY      1068 TTCGGGAGAGACTGTGATCTGTGCGCATCCCTTTTCTTGGGGTGGGCTTACGTTT 1127
         |||||
Db      100738 TTCGGGAGAGACTGTGATCTGTGCGCATCCCTTTTCTTGGGGTGGGCTTACGTTT 100797

QY      1128 GCCAAGCTGCGCTTAAGCCCATATGTACAAATGCGGTTCAGGAAGAAATGAGG 1187
         |||||
Db      100798 GCCAAGCTGCGCTTAAGCCCATATGTACAAATGCGGTTCAGGAAGAAATGAGG 100857
  
```

Oy	1188	AAGATTTTTCGTGCTCTTTTTCACAGAAAGAGGCCATTTTTACAGTAGCTGTC	1247
Db	100858	AAGATTTTTCGTGCTCTTTTTCACAGAAAGAGGCCATTTTTACAGTAGCTGTC	100917
Oy	1248	AGCGAATAAGACTTGTCGTATTATTTTCAGCTTAAGCTCTGGTGCAGGTGAACACG	1307
Db	100918	AGCGAATAAGACTTGTCGTATTATTTTCAGCTTAAGCTCTGGTGCAGGTGAACACG	100977
Oy	1308	GTCGCACTGTAAAGGAGCTTAAGCTTAAGAAAAGCCACCAGCTGGCCTGCTTTAAAA	1367
Db	100978	GTCGCACTGTAAAGGAGCTTAAGCTTAAGAAAAGCCACCAGCTGGCCTGCTTTAAAA	101037
Oy	1368	TACCCGACTTCCAAGACGAGCATCTAACGGAGCCAGCAAATTAAGAAATGATCGCTCAGT	1427
Db	101038	TACCCGACTTCCAAGACGAGCATCTAACGGAGCCAGCAAATTAAGAAATGATCGCTCAGT	101097
Oy	1428	ATAAAAATATTTTTCTTAAAGAAGCTTCTTAATGGCTCTTTTGTAACTTTT	1483
Db	101098	ATAAAAATATTTTTCTTAAAGAAGCTTCTTAATGGCTCTTTTGTAACTTTT	101153

RESULT 9	AR228218	181 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR228218				
DEFINITION	Sequence 14 from patent US 6448005.				
ACCESSION	AR228218				
VERSION	AR228218.1	GI:27266965			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 181)				
AUTHORS	Glucksmann,M.A. and Tsai,F.-Y.				
TITLE	14723 Receptor, a novel G-protein coupled receptor				
JOURNAL	Patent: US 6448005-A 14 10-SEP-2002;				
FEATURES	Location/Qualifiers				
SOURCE	1..181				
	/organism="unknown"				
	/mol_type="genomic DNA"				

ORIGIN	
Query Match	11.6%; Score 181; DB 6; Length 181;
Best Local Similarity	100.0%; Pred. No. 2.7e-94;
Matches 181; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

Oy	1021	TGTGGAGTCCCATCATCATCACCATCTCTCATCTTGAATCCAAAATCCGGCAGACG	1080
Db	1	TGTGGAGTCCCATCATCATCACCATCTCTCATCTTGAATCCAAAATCCGGCAGACG	60
Oy	1081	TGTCATCTGGGCATCCCTTTTCTTCTGGGTGGTGGCCTTCAAGITTTGCCAATCTGCCC	1140
Db	61	TGTCATCTGGGCATCCCTTTTCTTCTGGGTGGTGGCCTTCAAGITTTGCCAATCTGCCC	120
Oy	1141	TAAACCCCATCTGTACAACAATGTGGCTTTCAAGAAAGAAATGAGAGAAATTTTGGCT	1200
Db	121	TAAACCCCATCTGTACAACAATGTGGCTTTCAAGAAAGAAATGAGAGAAATTTTGGCT	180
Oy	1201	G 1201	
Db	181	G 181	

RESULT 10	AR228219	138 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR228219				
DEFINITION	Sequence 15 from patent US 6448005.				
ACCESSION	AR228219				
VERSION	AR228219.1	GI:27266966			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 138)				

ORIGIN	Query Match	8.8%; Score 138; DB 6; Length 138;
Beat Local Similarity	100.0%; Pred. No. 4,9e-69;	
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
FEATURES		
source	1..138	
	/organism="unknown"	
	/mol_type="genomic DNA"	
Db	1 AAGAGGCTTACGCTGAGCTTGGCACTCTGAGAGCCACGATCCGAGTGTCCCAACAA	60
Qy	969 GACACACGACCTCTCCGACGCTCTTCCTGCTCAGGTTTCCTTCATCATGTGAGGT	1028
Db	61 GACACACGACCTCTCCGACGCTCTTCCTGCTCAGGTTTCCTTCATCATGTGAGGT	120
Qy	1029 CCATCATCATCACCATC 1046	
Db	121 CCATCATCATCACCATC 138	
RESULT 11		
LOCUS	BC049237	835 bp
DEFINITION	Mus musculus electron transferring flavoprotein, beta polypeptide,	linear ROD 07-OCT-2003
ACCESSION	BC049237	
VERSION	BC049237.1	GI:29351580
KEYWORDS	MGC.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	1 (bases 1 to 835)	
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.M.,	
	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,	
	Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhac,N.K.,	
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heien,F.,	
	Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,	
	Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,	
	Scheetz,T.E., Brownstein,M.J., Ueda,T.B., Toshiyuki,S.,	
	Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,	
	Abraham,R.D., Mulvaney,S.J., Bosak,S.A., McEwan,P.J.,	
	McEranon,K.U., Malek,U.A., Gunaratne,P.H., Richards,S.,	
	Wolley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,	
	Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,	
	Fahney,J., Hellon,E., Kettman,M., Madan,A., Young,A.C., Rodriguez,S.,	
	Sanchez,A., Whiting,M., Madan,A., Madan,A.C., Shevchenko,Y.,	
	Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,	
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,	
	Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smallov,D.E.,	
	Scherech,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
	Generation and initial analysis of more than 15,000 full-length	
	human and mouse cDNA sequences	
	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
TITLE	22388257	
JOURNAL	12477932	
MEDLINE	2 (bases 1 to 835)	
PUBMED	Strausberg,R.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (21-MAR-2003) National Institutes of Health, Mammalian	
TITLE	Gene Collection (MGC), Cancer Genomes Office, National Cancer	
JOURNAL	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
	USA	
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov	
COMMENT	Contact: MGC help desk	
	Email: ggapbs-romail.nih.gov	
	Tissue Procurement: Jeffrey E. Green, M.D.	

NEIGHBORING SEQUENCE INFORMATION: This sequence is the entire insert of the clone.

FEATURES
Source
Location/Qualifiers
1. 169206
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/map="7"
/clone="RP24-289C13"
/clone_1id="RPCL-24"
1. .90
repeat_region
/rpt_family="B2"
398.548
/rpt_family="Alu"
repeat_region
1705.1784
/rpt_family="MIR"
repeat_region
2903.3072
/rpt_family="B4"
repeat_region
3609.4021
/rpt_family="MaLR"
repeat_region
6106.6248
/rpt_family="Alu"
repeat_region
7008.7052
/rpt_family="ERV1"
repeat_region
7053.7205
/rpt_family="B4"
repeat_region
7206.7343
/rpt_family="ERV1"
repeat_region
7504.7773
/rpt_family="L1"
repeat_region
7804.8139
/rpt_family="MaLR"
repeat_region
8217.8597
/rpt_family="MaLR"
repeat_region
8598.8652
/rpt_family="L1"
repeat_region
8621.9123
/rpt_family="L1"
repeat_region
9159.9275
/rpt_family="L1"
repeat_region
9288.9434
/rpt_family="Alu"
repeat_region
9454.9577
/rpt_family="L1"
repeat_region
9648.9807
/rpt_family="L1"
repeat_region
9813.9956
/rpt_family="Alu"
repeat_region
9958.10577
/rpt_family="L1"
repeat_region
11055.11618
/rpt_family="L1"
repeat_region
11613.11840
/rpt_family="B2"
repeat_region
11841.12020
/rpt_family="L1"
repeat_region
13831.13978
/rpt_family="L1"
repeat_region
13879.14312
/rpt_family="MaLR"
repeat_region
14317.14730
/rpt_family="L1"
repeat_region
14731.14915
/rpt_family="B2"
repeat_region
14916.14967
/rpt_family="L1"
repeat_region
16136.16585
/rpt_family="L1"
repeat_region
17108.18100
/rpt_family="L1"
unSURE
18285.18307

/note="Sequence derived from one plasmid subclone."
repeat_region
18359.18501
/rpt_family="Alu"
repeat_region
18511.19194
/rpt_family="L1"
repeat_region
19246.19380
/rpt_family="B2"
repeat_region
19381.19425
/rpt_family="B4"
repeat_region
19600.19790
/rpt_family="B2"
repeat_region
19869.20088
/rpt_family="B4"
repeat_region
20115.20270
/rpt_family="B4"
repeat_region
20581.20778
/rpt_family="L1"
repeat_region
21094.21186
/rpt_family="L1"
repeat_region
21228.21307
/rpt_family="MaLR"
repeat_region
21301.21491
/rpt_family="B2"
repeat_region
21499.21743
/rpt_family="MaLR"
repeat_region
21822.22007
/rpt_family="L1"
repeat_region
22126.22216
/rpt_family="B4"
repeat_region
23839.24140
/rpt_family="L1"
repeat_region
25383.25460
/rpt_family="ID"
repeat_region
25388.25456
/product="tRNA-Ala"
/note="Likely pseudogene (HMM Sc=34.12 / Sec struct Sc=1.92)"
repeat_region
26997.27316
/rpt_family="ERV1"
repeat_region
27476.27572
/rpt_family="Alu"
repeat_region
28060.28287
/rpt_family="L2"
repeat_region
28441.28607
/rpt_family="MaLR"
repeat_region
28752.28828
/rpt_family="MaLR"
repeat_region
28829.29172
/rpt_family="MaLR"
repeat_region
29173.29340
/rpt_family="MaLR"
repeat_region
29420.29503
/rpt_family="B4"
repeat_region
29504.29635
/rpt_family="Alu"

Query Match 5.3%; Score 92; DB 10; Length 169206;
Best Local Similarity 100.0%; Pred.No.1.6e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 AGCCTCTTCACATGCAATCTCACAGAGGGGTTTCATGAGATGCTTTCACACATCAGTGAC 77
DB 80518 AGCCTCTTCACATGCAATCTCACAGAGGGGTTTCATGAGATGCTTTCACACATCAGTGAC 80459
QY 78 CACTCCAGACTTGCAGGCTTTACCCGAACTCT 109
DB 80458 CACTCCAGACTTGTCCGACTTTACCCGAACTCT 80427

RESULT 13
AC079543 277603 bp DNA linear HTG 02-SEP-2000
LOCUS AC079543
DEFINITION Mus musculus clone RP23-381P16, WORKING DRAFT SEQUENCE, 67

unordered pieces.
ACCESSION AC079543 GI:9964908
VERSION AC079543.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 277603)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 277603)
DOE Joint Genome Institute.
Direct Submission
Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1883552
Center clone name: RPCI-23_381P16

Summary Statistics
Consensus quality: 224010 bases at least Q40
Consensus quality: 249280 bases at least Q30
Consensus quality: 255396 bases at least Q20
Estimated insert size: 271000; agarose-fp estimation
Estimated insert size: 271003; sum-of-contigs estimation
Quality coverage: 6.92 in Q20 bases; agarose-fp estimation
Quality coverage: 4.39 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1043: contig of 1043 bp in length
1044 1143: gap of unknown length
1144 2187: contig of 1044 bp in length
2188 2287: gap of unknown length
2288 3468: contig of 1181 bp in length
3469 3569: gap of unknown length
3569 4840: contig of 1272 bp in length
4841 4940: gap of unknown length
4941 6012: contig of 1072 bp in length
6013 6112: gap of unknown length
6113 7304: contig of 1192 bp in length
7305 7404: gap of unknown length
7405 8702: contig of 1298 bp in length
8703 8802: gap of unknown length
8803 10001: contig of 1199 bp in length
10002 10101: gap of unknown length
10102 11249: contig of 1148 bp in length
11250 11349: gap of unknown length
11350 12629: contig of 1280 bp in length
12630 12729: gap of unknown length
12730 12730: gap of unknown length
12730 13945: contig of 1216 bp in length
13946 14045: gap of unknown length
14046 15833: contig of 1788 bp in length
15834 15933: gap of unknown length
15934 17862: contig of 1929 bp in length
17863 17962: gap of unknown length
17963 18991: contig of 1029 bp in length
18992 19091: gap of unknown length
19092 20898: contig of 1807 bp in length
20899 20998: gap of unknown length
20999 22832: contig of 1834 bp in length

22833 22932: gap of unknown length
22933 22933 24144: contig of 1212 bp in length
24144 24244: gap of unknown length
24245 25669: contig of 1425 bp in length
25670 25769: gap of unknown length
25770 27239: contig of 1470 bp in length
27240 27339: gap of unknown length
27340 28790: contig of 1451 bp in length
28791 30106: contig of 1216 bp in length
28891 30206: gap of unknown length
30107 31602: contig of 1396 bp in length
30207 31603 31702: gap of unknown length
31603 32867: contig of 1165 bp in length
31703 32967: gap of unknown length
32868 34416: contig of 1449 bp in length
32968 34516: gap of unknown length
34417 35577: contig of 1061 bp in length
34517 35677: gap of unknown length
35578 37248: contig of 1571 bp in length
35678 37348: gap of unknown length
37249 38610: contig of 1262 bp in length
37349 38710: gap of unknown length
38611 40389: contig of 1679 bp in length
38711 40489: gap of unknown length
40390 41593: contig of 1104 bp in length
40490 41693: gap of unknown length
41594 43102: contig of 1409 bp in length
41694 43202: gap of unknown length
43103 44777: contig of 1575 bp in length
43203 44777: gap of unknown length
44778 44877: gap of unknown length
44878 46440: contig of 1563 bp in length
44879 46540: gap of unknown length
46441 48609: contig of 2069 bp in length
46541 48709: gap of unknown length
48710 49895: contig of 1186 bp in length
48711 49995: gap of unknown length
49896 51506: contig of 1511 bp in length
49996 51606: gap of unknown length
51507 53008: contig of 1402 bp in length
51607 53108: gap of unknown length
53009 55140: contig of 2032 bp in length
53109 55240: gap of unknown length
55141 57760: contig of 2520 bp in length
55241 57860: gap of unknown length
57761 59903: contig of 2043 bp in length
57861 60003: gap of unknown length
59904 62375: contig of 2372 bp in length
60004 62475: gap of unknown length
62376 64891: contig of 2416 bp in length
62476 64991: gap of unknown length
64892 67699: contig of 2708 bp in length
64992 67799: gap of unknown length
67700 69624: contig of 1825 bp in length
67800 69724: gap of unknown length
69625 71722: contig of 1998 bp in length
69725 71822: gap of unknown length
71723 71823 75072: contig of 3250 bp in length
71823 75172: gap of unknown length
75073 77256: contig of 2084 bp in length
75173 77356: gap of unknown length
77257 79802: contig of 2446 bp in length
77357 79902: gap of unknown length
79803 81721: contig of 1819 bp in length
79903 81722: gap of unknown length
81722 86203: contig of 4382 bp in length
81822 86303: gap of unknown length
86204 92047: contig of 5744 bp in length
86304 92147: gap of unknown length
92048 95857: contig of 3710 bp in length
92148 95957: gap of unknown length
95858 99755: contig of 3798 bp in length
95958 99855: gap of unknown length
99756

* consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 13733: contig of 13733 bp in length
 * 13734 13833: gap of unknown length
 * 13834 202133: contig of 188300 bp in length
 * 202134 202233: gap of unknown length
 * 202234 205371: contig of 3138 bp in length
 * 205372 205471: gap of unknown length
 * 205472 206631: contig of 1160 bp in length
 * 206632 206731: gap of unknown length
 * 206732 207869: contig of 1138 bp in length
 * 207870 207969: gap of unknown length
 * 207970 209084: contig of 1115 bp in length
 * 209085 209184: gap of unknown length
 * 209185 210652: contig of 1468 bp in length
 * 210653 210753: gap of unknown length
 * 210753 212822: contig of 2070 bp in length.

FEATURES

source

1. 212822
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="RP32-379B20"
 1. 1448
 /note="wgs_contig"
 13834. 14900
 /note="wgs_contig"
 198192. 159489
 /note="wgs_contig"
 200918. 202133
 /note="wgs_contig"

ORIGIN

Query Match 4.9%; Score 77; DB 2; Length 212822;
 Best Local Similarity 100.0%; Pred. No. 1.1e-32;
 Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 987 ACGCTCTTCCTGCTCATGTTCTTCATCATGAGATCCATCATCATC 1046
 |||||
 Db 114483 ACGCTCTTCCTGCTCATGTTCTTCATCATGAGATCCATCATCATC 114542
 |||||
 Qy 1047 CTCCTCATCTTGATCCA 1063
 |||||
 Db 114543 CTCCTCATCTTGATCCA 114559

RESULT 15
 AC107608 225030 bp DNA linear ROD 30-JUL-2003
 LOCUS Rattus norvegicus 1 BAC CH230-204P6 (Children's Hospital Oakland
 DEFINITION Research Institute) complete sequence.
 AC107608
 VERSION AC107608.6 GI:33342287
 KEYWORDS HTG.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 225030)
 Wuzny,D,Marie, Metzker,M,lee, Abramzon,S, Adams,C, Alder,J.,
 Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D.,
 Ayalebechi,V, Ayogbi,A, Ayodeji,M, Baca,E, Baden,H.,
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F.,
 Bielawski,K, Blair,U, Blankenburg,K, Blyth,P, Brown,M.,
 Bryant,N, Burch,C, Burch,P, Burrell,K, Calderon,E,
 Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Garcia,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
 Hollins,B., Howell,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kover,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensunewa,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartine,M., Mahmoud,M., Mallory,K., Mangum,A.,
 Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,B.,
 Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
 Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
 Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
 Newton,N., Nguyen,N., Norris,S., Nwachileme,O., Okwono,G.,
 Olampunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
 Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poindexter,A.,
 Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quirroz,J., Rachlin,E.,
 Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
 Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
 Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Savary,G., Scherer,S.,
 Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
 Sisson,I., Sitter,C.D., Smajs,D., Snead,A., Sodestgen,E.,
 Song,X.-Z., Sorelle,R., Sosa,J., Steidle,M., Strong,R., Sutton,A.,
 Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,D.,
 Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseña,D.,
 Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
 Warren,R., Wei,X., White,F., Williams,G., Wilson,R., Wlecyk,R.,
 Woodson,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
 Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
 Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
 Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

```

repeat_region      837. .955
                    /rpt_family="L3"
repeat_region      2039. .2146
                    /rpt_family="(TG)n"
repeat_region      complement(2936. .3141)
                    /rpt_family="B3"
repeat_region      3253. .3289
                    /rpt_family="(TATAA)n"
repeat_region      3607. .3653
                    /rpt_family="(TG)n"
repeat_region      complement(3757. .3883)
                    /rpt_family="B1F"
repeat_region      complement(4564. .4662)
                    /rpt_family="ID_Rn"
repeat_region      complement(4814. .4859)
                    /rpt_family="L2"
repeat_region      5719. .5849
                    /rpt_family="MIR"
repeat_region      complement(6019. .6120)
                    /rpt_family="PB1D10"
repeat_region      6385. .6429
                    /rpt_family="(TTA)n"
repeat_region      6697. .6780
                    /rpt_family="AT_rich"
repeat_region      7022. .7066
                    /rpt_family="(TG)n"
repeat_region      7121. .7164
                    /rpt_family="(TG)n"
repeat_region      7227. .7278
                    /rpt_family="(TG)n"
repeat_region      7341. .7370
                    /rpt_family="(TG)n"
repeat_region      7437. .7462
                    /rpt_family="(TG)n"
repeat_region      7558. .7635
                    /rpt_family="(TG)n"
repeat_region      8060. .8310
                    /rpt_family="CT-rich"
repeat_region      8330. .8387
                    /rpt_family="L2"
repeat_region      8511. .8870
                    /rpt_family="(CA)n"
repeat_region      9441. .9608
                    /rpt_family="MLT1A0"
repeat_region      9622. .9956
                    /rpt_family="URRAA"
repeat_region      10552. .10588
                    /rpt_family="(TG)n"
repeat_region      10589. .10621
                    /rpt_family="(GA)n"
repeat_region      11228. .11701
                    /rpt_family="L1_Rac3"
repeat_region      complement(11885. .12186)
                    /rpt_family="B4"
repeat_region      12257. .12329
                    /rpt_family="RSINE1"
repeat_region      13079. .13130
                    /rpt_family="RSINE1"
repeat_region      13967. .14097
                    /rpt_family="ID_B1"
repeat_region      15038. .15578
                    /rpt_family="Lx2"
repeat_region      complement(15802. .15961)
                    /rpt_family="MON1"
repeat_region      16995. .17099
                    /rpt_family="B1F"
repeat_region      complement(17242. .17382)
                    /rpt_family="B1_Mur2"
repeat_region      17516. .17578
                    /rpt_family="CT-rich"
repeat_region      complement(17792. .17867)
                    /rpt_family="LIMC4"
repeat_region      17907. .18012

```

```

repeat_region      /rpt_family="ID_B1"
                    17927. .18026
repeat_region      /rpt_family="B1F1"
                    18028. .18066
repeat_region      /rpt_family="(TCTG)n"
                    18167. .18237
repeat_region      /rpt_family="ID4_"
                    20813. .20858
repeat_region      /rpt_family="(CAA)n"
                    21044. .21159
repeat_region      /rpt_family="B4"
                    21172. .21278
repeat_region      /rpt_family="(CA)n"
                    complement(21619. .22012)
repeat_region      /rpt_family="MLT1C"
                    22038. .22317
repeat_region      /rpt_family="L2"
                    complement(22432. .22509)
repeat_region      /rpt_family="ID4_"
                    22724. .22799
repeat_region      /rpt_family="B1F"
                    22967. .23021
repeat_region      /rpt_family="CT-rich"
                    23905. .24116
repeat_region      /rpt_family="B2_Rn"
                    complement(24255. .24400)
repeat_region      /rpt_family="ID_B1"
                    24284. .24491
repeat_region      /rpt_family="B1"
                    24409. .24445
repeat_region      /rpt_family="(TTG)n"
                    24714. .24748
repeat_region      /rpt_family="(CA)n"
                    25777. .25806

```

```

Query Match      4.9%; Score 77; DB 10; Length 225030;
Best Local Similarity 100.0%; Pred. No. 1,1e-32;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      987  ACCGCTTCTGCTCATGATGTTTCTTTCATCATGATGAGTCCATCATCATCATC 1046
DB      142982  ACCGCTTCTGCTCATGATGTTTCTTTCATCATGATGAGTCCATCATCATCATC 143041

QY      1047  CTCCTCATCTTGATCCA 1063
DB      143042  CTCCTCATCTTGATCCA 143058

```

Search completed: October 1, 2004, 01:42:01
Job time : 5927.2 secs

This Page Blank (uspio)



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 19:01:32 ; Search time 569.119 Seconds
(without alignments)
1164.646 Million cell updates/sec

Title: US-10-077-698-5

Sequence: 1 ttcgccagctcgcgtcagc.....aaaaaaaaaaaaaaaa 1560

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018a:*
5: geneseqn20018b:*
6: geneseqn20028:*
7: geneseqn20038a:*
8: geneseqn20038b:*
9: geneseqn20038c:*
10: geneseqn20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1560	100.0	1560	3	AAA64347
2	1560	100.0	1560	3	ABQ81227 Mouse 142
3	1212	77.7	1561	3	AAZ49746 Murine 14
4	180	11.5	180	3	AAA64344
5	138	8.8	138	3	AAA64345
6	44	2.8	1086	6	ABT04869 Human G P
7	44	2.8	1086	6	ABST73398 DNA encod
8	44	2.8	1086	6	ABST73399 DNA encod
9	44	2.8	1086	6	ABST73343 CDNA enco
10	44	2.8	1086	7	ACA60998 CDNA enco
11	44	2.8	1086	8	ADB47641 Human CDN
12	44	2.8	1160	7	ABX10627 DNA encod
13	44	2.8	1321	4	AAD08854 Human G-P
14	44	2.8	1458	4	AAI66039 Human G P
15	44	2.8	1737	7	ABZ42599 Human G P
16	44	2.8	1742	3	AAZ49745 Human 142
17	44	2.8	1743	3	AAA64346 DNA encod
18	44	2.8	1743	6	ABQ81226 Human 142
19	44	2.7	933	9	ADC12755 Human GPC
20	40	2.6	1066	9	ADC87012 Human GPC
21	40	2.6	1104	4	AAZ42815 Human G P
22	40	2.6	3173	4	AAZ42815 Human G P
23	39	2.5	638	6	AAZ42815 Human G P

24	38	2.4	241	3	AAZ80388 Human col
25	38	2.4	459	8	ACH44239 Human foe
26	38	2.4	472	8	ACH40136 Human foe
27	38	2.4	487	3	ACH02281 Human foe
28	38	2.4	489	8	ACH45264 Human foe
29	38	2.4	490	8	ACH13024 Human adu
30	38	2.4	500	5	ABV60845 Human pro
31	38	2.4	517	8	ACH43949 Human foe
32	38	2.4	762	2	AAZ39905 Gastric c
33	38	2.4	768	6	ABL39413 Human ele
34	38	2.4	813	2	AAZ39904 Gastric c
35	38	2.4	912	5	AAZ70364 DNA encod
36	38	2.4	919	8	AAZ68786 Human pan
37	38	2.4	930	5	AAZ70067 DNA encod
38	38	2.4	1281	7	ADA52753 Human cod
39	38	2.4	2631	5	AAZ64404 DNA encod
40	38	2.4	29222	6	ABL39412 Human ele
41	29	1.9	469	8	ACH45354 Human foe
42	28	1.8	445	8	ACH34462 Human end
43	28	1.8	459	8	ACH28569 Human adu
44	28	1.8	1291	3	AAZ67666 Human sec
45	27	1.7	27	6	ABQ81230 Mouse 142

ALIGNMENTS

RESULT 1	AAA64347	standard; DNA; 1560 BP.
XX	AAA64347	
AC	AAA64347	
DT	20-DEC-2000	(first entry)
XX		
DE	DNA encoding a murine G-protein coupled receptor designated 14273.	
XX		
KV	Mouse; G-protein coupled receptor; receptor-mediated disorder;	
KW	14273 receptor; cardiovascular disease; congestive heart failure;	
KM	cardiac myocyte hypertrophy; ss.	
XX		
OS	Mus sp.	
XX		
FN	Key	Location/Qualifiers
FT	CDS	195..1280
FT	FT	/*tag= a
XX		/product= "G-protein coupled receptor 14273"
XX		
XX	W0200050596-A2.	
XX		
PD	31-AUG-2000.	
XX		
PF	28-FEB-2000; 2000MO-US005068.	
XX		
PR	26-FEB-1999; 99US-00261599.	
PR	08-DEC-1999; 99US-00456455.	
XX		
PA	(MILL-) MILLENIUM PHARM INC.	
XX		
PI	Gluckemann MA, Tsai F;	
XX		
DR	WPI; 2000-587184/55.	
DR	P-PSDB; AAB08539.	
XX		
PT	14273 receptor polynucleotides and polypeptides, useful in the diagnosis	
PT	and treatment of receptor-mediated disorders such as cardiovascular	
PT	diseases.	
XX		
PS	Claim 3; Fig 7; 105pp; English.	
XX		
CC	The present sequence encodes a murine G-protein coupled receptor,	
CC	designated 14273 receptor. The G-protein coupled receptor 14273	
CC	polypeptide is used to produce antibodies, in drug screening assays, and	

CC to screen for agonist and antagonists. The antibodies are used for
CC diagnostic applications, and to assess abnormal tissue distributions or
CC abnormal expression during development. The antibodies are also useful
CC for tissue typing, and in forensic identification, as well as for
CC inhibiting receptor function. The 14273 polynucleotides can be used to
CC express the protein, to detect 14273 mRNA, to detect genetic alterations
CC in the 14273 gene, and as a source of primers and probes. The 14273
CC polynucleotides and polypeptides are also useful as a target for
CC diagnosis and treatment of receptor-mediated disorders, especially
CC cardiovascular diseases such as congestive heart failure caused by
CC cardiac myocyte hypertrophy
XX

Sequence 1560 BP; 321 A; 449 C; 390 G; 400 T; 0 U; 0 Other;

Query Match 100.0%; Score 1560; DB 3; Length 1560;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTGGCAAGCTCAGCGTAAGCTCTTCCACTGCAATCTCAGAGAGGGGTTTCATGAGTGC 60
DB 1 TTGGCAAGCTCAGCGTAAGCTCTTCCACTGCAATCTCAGAGAGGGGTTTCATGAGTGC 60
QY 61 TTCACACCATCAGTACGACCACTCCAGACTTGTCCGGCTTTAACCGAATCTTCAACGCGAG 120
DB 61 TTCACACCATCAGTACGACCACTCCAGACTTGTCCGGCTTTAACCGAATCTTCAACGCGAG 120
QY 121 TTGATGACCCCTTGAAGAGCAGAGGGCGCGAGCTCCGCACTTCCCGAGGCGGTGG 180
DB 121 TTGATGACCCCTTGAAGAGCAGAGGGCGCGAGCTCCGCACTTCCCGAGGCGGTGG 180
QY 181 GCCGGGGCCCGGAGATGCCCTGAGTGTGACAGAGAGGCGCTGTGCTCCGACCA 240
DB 181 GCCGGGGCCCGGAGATGCCCTGAGTGTGACAGAGAGGCGCTGTGCTCCGACCA 240
QY 241 CCTTGAACCAAGTAATCCGACCCACTTCCCTTTCTTCTGAGATGTCAAGGCGACCA 300
DB 241 CCTTGAACCAAGTAATCCGACCCACTTCCCTTTCTTCTGAGATGTCAAGGCGACCA 300
QY 301 GGTGTGATGAGAGGTGTGAGAGCAACCGTTTGGGACATCATTTTGTGCTCAGTGC 360
DB 301 GGTGTGATGAGAGGTGTGAGAGCAACCGTTTGGGACATCATTTTGTGCTCAGTGC 360
QY 361 TGGGCAACGTTGTGCTTACTGTGTGTGAGCGCGCGCTGGGCGCTGAGGCGTCA 420
DB 361 TGGGCAACGTTGTGCTTACTGTGTGTGAGCGCGCGCTGGGCGCTGAGGCGTCA 420
QY 421 GCTGTGTGTCAACCTTCTTGTGCGCGGATTTGCTTTCACAGGCGCATCCCTTA 480
DB 421 GCTGTGTGTCAACCTTCTTGTGCGCGGATTTGCTTTCACAGGCGCATCCCTTA 480
QY 481 TCGTGTGCGCTGAGCTGAGGCGTGGCGTTGGGAGCGCGTGTGCAACCTGCTTCT 540
DB 481 TCGTGTGCGCTGAGCTGAGGCGTGGCGTTGGGAGCGCGTGTGCAACCTGCTTCT 540
QY 541 ACGTATGACAAATGAGGCGGAGCGTCAAGATCTTCAACTGCGCGGTCAGCTGAG 600
DB 541 ACGTATGACAAATGAGGCGGAGCGTCAAGATCTTCAACTGCGCGGTCAGCTGAG 600
QY 601 GCATGTGTGTGATGTGCGCTCCGGCGCGGCTTGAAGCGCGCGGCGCGGAGCT 660
DB 601 GCATGTGTGTGATGTGCGCTCCGGCGCGGCTTGAAGCGCGCGGCGCGGAGCT 660
QY 661 CGGACATGCTGCTTATATATGAGGTTACTGCGGCGTCCGCGGCTGCCCTCTA 720
DB 661 CGGACATGCTGCTTATATATGAGGTTACTGCGGCGTCCGCGGCTGCCCTCTA 720
QY 721 TGTTCGCGTGTGTCGCGAGCGCTTCCGCGGCGGAGCAGAGAAATTCGATTTG 780
DB 721 TGTTCGCGTGTGTCGCGAGCGCTTCCGCGGCGGAGCAGAGAAATTCGATTTG 780
QY 781 TGGATTTGGCCCAACCGCATAGAGAAATCTCATGAGATGTGTTTTTGAAGCT 840
DB 781 TGGATTTGGCCCAACCGCATAGAGAAATCTCATGAGATGTGTTTTTGAAGCT 840

```

```

QY 841 TCCGAGCCGGAGTGTGATGTGATCAGTTACTCCAAAATTTTACAGATCAGAG 900
DB 841 TCCGAGCCGGAGTGTGATGTGATCAGTTACTCCAAAATTTTACAGATCAGAG 900
QY 901 CATGCGGAAGAGGCTTACCTGAGCTTGGCATCTCTGAGGCGACAGATCCGAGT 960
DB 901 CATGCGGAAGAGGCTTACCTGAGCTTGGCATCTCTGAGGCGACAGATCCGAGT 960
QY 961 CCCAACAAGACTACCGACTCTTCGAGCGCTTCCGCTCATGAGTTTCTTCTATCA 1020
DB 961 CCCAACAAGACTACCGACTCTTCGAGCGCTTCCGCTCATGAGTTTCTTCTATCA 1020
QY 1021 TGTGAGTCCCATCATCATCATCATCATCTCTCATCTTGAATCAAAATTCGAGAG 1080
DB 1021 TGTGAGTCCCATCATCATCATCATCATCTCTCATCTTGAATCAAAATTCGAGAG 1080
QY 1081 TGTGATCTGAGCATCCCTTTCTTCTGAGTGTGAGCTTCAAGTTCGCACTG 1140
DB 1081 TGTGATCTGAGCATCCCTTTCTTCTGAGTGTGAGCTTCAAGTTCGCACTG 1140
QY 1141 TAAACCCCATATCTGTAACAATGTCGTTCAAGAAACGAATGAGAGAAATTTT 1200
DB 1141 TAAACCCCATATCTGTAACAATGTCGTTCAAGAAACGAATGAGAGAAATTTT 1200
QY 1201 GCTTCTTTTTCAGAGAAAGGAGCCATTTTACAGATAGCTGTGAGCGAATG 1260
DB 1201 GCTTCTTTTTCAGAGAAAGGAGCCATTTTACAGATAGCTGTGAGCGAATG 1260
QY 1261 TGTCTGTTATTTTCAAGTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
DB 1261 TGTCTGTTATTTTCAAGTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
QY 1321 GGGAGTTAATCTTAAGAAAGCCACAGTGGCGCTTTAAATATACCGGACTTCA 1380
DB 1321 GGGAGTTAATCTTAAGAAAGCCACAGTGGCGCTTTAAATATACCGGACTTCA 1380
QY 1381 ACAGAGGACATCTACGAGGCGACGAAATTAAGAAATGATGCTCAGTATTAAT 1440
DB 1381 ACAGAGGACATCTACGAGGCGACGAAATTAAGAAATGATGCTCAGTATTAAT 1440
QY 1441 TCCTTAAAGAACTTCTATGAGTGTCTTGTGAACTTTTAAAGTGTGTAAT 1500
DB 1441 TCCTTAAAGAACTTCTATGAGTGTCTTGTGAACTTTTAAAGTGTGTAAT 1500
QY 1501 GATCTAGTTAATTAATTTTATTAAGTGTCTTCAAAAAAATTTTAAAAA 1560
DB 1501 GATCTAGTTAATTAATTTTATTAAGTGTCTTCAAAAAAATTTTAAAAA 1560

```

RESULT 2
AB081227
ID AB081227 standard; cDNA; 1560 BP.

XX AB081227;

DT 05-DEC-2002 (first entry)

XX Mouse 14273 nucleic acid, associated with metabolic disorder.

XX Mouse; 14273; metabolic disorder; obesity; diabetes; anorexia; cachexia;
XX anorectic; antidiabetic; anabolic; transgenic animal; gene therapy; gene;
XX ss.

OS Mus musculus.

XX Key Location/Qualifiers

XX Key 195..1280

XX CDS /*tag= a

XX FT /product= "14273"

XX W0200267868-A2.

PD 06-SEP-2002.
 XX
 PF 26-FEB-2002; 2002W0-US006131.
 XX
 PR 26-FEB-2001; 2001US-0271655P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Gimeno R, Tsai F;
 XX
 DR WPI; 2002-698629/75.
 DR P-PSDB; ABB79907.
 XX
 PT Identifying a nucleic acid associated with a metabolic disorder, useful
 PT for diagnosing metabolic disorders, e.g. obesity, comprises contacting
 PT the sample with a probe comprising at least 25 contiguous nucleotides of
 PT the 14273 gene.
 XX
 PS Claim 1; Fig 2; 95pp; English.
 XX
 CC The present sequence is that of a murine nucleic acid, designated 14273,
 CC associated with metabolic disorders. 14273 molecules are expressed at
 CC high levels in adipose tissue, e.g. white adipose tissue and brown
 CC adipose tissue, as well as in pancreatic islets. They are upregulated
 CC during exposure to cold (i.e. under conditions that affect brown or white
 CC adipocyte metabolism) and downregulated in genetic models of obesity.
 CC 14273 knock-out mice, when fed a high-fat diet, gain more weight and have
 CC larger epididymal fat pads than wild-type mice. They also show increased
 CC levels of glucose and insulin upon fasting. A 14273 agonist may be
 CC beneficial to the treatment of obesity and/or type II diabetes by
 CC preventing fat accumulation on a high-fat diet and/or the increases in
 CC endogenous glucose production which occur in type II diabetes. The
 CC present invention provides 14273 nucleic acids, polypeptides and
 CC antibodies useful for the diagnosis and treatment of metabolic disorders
 CC including obesity, anorexia, cachexia and diabetes. Also provided are
 CC methods for identifying a subject having a metabolic disorder, for
 CC identifying a compound capable of modulating metabolic activity, for
 CC for modulating metabolic activity or adipocyte activity (hyperplastic
 CC growth, hypertrophic growth or lipogenesis), methods for modulating
 CC endogenous glucose levels
 CC
 SO Sequence 1560 BP; 321 A; 449 C; 390 G; 400 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1560; DB 6; Length 1560;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	361	TGGGCAACGCTGTGTCTCTAGTGTGTGGGCGCCCGCTGGCGCGCTGGGCGCTCAAGCA	420
Qy	421	GCTGTGTCTCAACCTCTTTCGCGGGAATTTGCTTTACACAGCGCATCCCTTCTAGTC	480
Db	421	GCTGTGTCTCAACCTCTTTCGCGGGAATTTGCTTTACACAGCGCATCCCTTCTAGTC	480
Qy	481	TGCTGT	540
Db	481	TGCTGT	540
Qy	541	ACGTGATGACAAATAGGCGGAGCGTCACAGTCTCACTGCGCGGCTGAGCTGGAGC	600
Db	541	ACGTGATGACAAATAGGCGGAGCGTCACAGTCTCACTGCGCGGCTGAGCTGGAGC	600
Qy	601	GATGT	660
Db	601	GATGT	660
Qy	661	CGGCACTGT	720
Db	661	CGGCACTGT	720
Qy	721	TGTTCCGCGT	780
Db	721	TGTTCCGCGT	780
Qy	781	TGATTTGCGCCCAACCGGATAGAGAAATCTCATGGAGTGTGTGTGTGTGTGTGTGT	840
Db	781	TGATTTGCGCCCAACCGGATAGAGAAATCTCATGGAGTGTGTGTGTGTGTGTGTGT	840
Qy	841	TCTGT	900
Db	841	TCTGT	900
Qy	901	CATGCGGGAAGAGGCTTACGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Db	901	CATGCGGGAAGAGGCTTACGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Qy	961	CCCAACAAAGCTACGAGCTTTCGCGAGGCTTCTGCTCATGTGTGTGTGTGTGTGT	1020
Db	961	CCCAACAAAGCTACGAGCTTTCGCGAGGCTTCTGCTCATGTGTGTGTGTGTGTGT	1020
Qy	1021	TGTGAGATCCCAT	1080
Db	1021	TGTGAGATCCCAT	1080
Qy	1081	TGTCATCTGGGCAATCCCTTTCTTCTGCGGAGGAGGCTTCAAGTTGGCAATCTGCC	1140
Db	1081	TGTCATCTGGGCAATCCCTTTCTTCTGCGGAGGAGGCTTCAAGTTGGCAATCTGCC	1140
Qy	1141	TAAACCCCATATCTGTATACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1200
Db	1141	TAAACCCCATATCTGTATACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1200
Qy	1201	GCTTCTTTTTCAGAGAAAGGAGCCATTTTACAGATACGTGTGTGTGTGTGTGTGTGT	1260
Db	1201	GCTTCTTTTTCAGAGAAAGGAGCCATTTTACAGATACGTGTGTGTGTGTGTGTGTGT	1260
Qy	1261	TGTCTGTATTTTCCAGTAACTAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1320
Db	1261	TGTCTGTATTTTCCAGTAACTAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1320
Qy	1321	GGAAGTTAACTTCAAGAAAGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1380
Db	1321	GGAAGTTAACTTCAAGAAAGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1380
Qy	1381	ACAGCAGGCACTTACAGAGCAGCAATTAAGAAATGATGTGTGTGTGTGTGTGTGTGT	1440
Db	1381	ACAGCAGGCACTTACAGAGCAGCAATTAAGAAATGATGTGTGTGTGTGTGTGTGTGT	1440
Qy	1441	TTCCTTAAAGAACTTTTATAGGTTCTTTTGTGAATTTTAAAGTGTGTGTGTGTGT	1500
Db	1441	TTCCTTAAAGAACTTTTATAGGTTCTTTTGTGAATTTTAAAGTGTGTGTGTGTGT	1500

Dd	1441	TCCCTTAAAGAACCTTCATGCGGTTCCTTTGGTAACCTTTTTTAAGTGCTGTTTGAATAAT	1560
Oy	1501	GATCAGTAGTAATTAATTTTATTATTATTAACGCGTTCTCAACAAAAA	1560
Dd	1501	GATCAGTAGTAATTAATTTTATTATTATTAACGCGTTCTCAACAAAAA	1560
<hr/>			
RESULT 3			
ID	AAZ49746	standard; cDNA, 1561 BP.	
XX	AAZ49746;		
DT	02-MAY-2000	(first entry)	
XX	Murine 14273 G-protein coupled receptor (GPCR) encoding cDNA.		
KW	G-protein coupled receptor; GPCR; 14273 receptor; mouse; somatostatin;		
KW	cellular function/activity; galanin receptor; chemokine receptor;		
KW	diagnosis; treatment; receptor-mediated disorder; screening; ss.		
XX	Mus sp.		
OS			
FH	Key	Location/Qualifiers	
FT	CDS	195..1280	
FT		/tag= a	
FT		/product= "Murine 14273 G-protein coupled receptor"	
FT		/note= "Ortholog of human 14273 GPCR"	
FT		/transl_except= (pos:228..230, aa:Gly)	
FT		/transl_except= (pos:315..317, aa:Val)	
FT		/transl_except= (pos:1002..1004, aa:Ile)	
FT	unSURE	987..989	
FT		/*tag= b	
FT		/note= "Encodes Thr"	
XX			
PN	WO200000611-A2.		
PD			
PP	06-JAN-2000.		
PR	30-JUN-1999;	99WO-US014842.	
XX			
PR	30-JUN-1998;	98US-.00107761.	
PR	30-DEC-1998;	98US-.00223538.	
PR	26-FEB-1999;	99US-.00261599.	
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
PI	Gluckesmann MA, Tsai F;		
DR	WPI; 2000-147268/13.		
DR	P-PsDB; AAV44663.		
PT	Novel G-protein coupled receptor 14273 polynucleotides and polypeptides		
PT	used as a target for diagnosis and treatment of receptor-mediated		
PT	disorders.		
XX			
PS	Claim 3a; Fig 7; 89pp; English.		
XX			
CC	The present sequence is the cDNA encoding mouse G-protein coupled		
CC	receptor (GPCR) 14273 protein, an ortholog of human 14273 GPCR. It is		
CC	involved in modulation of a cellular function or activity upon binding of		
CC	a ligand to the GPCR. The GPCR 14273 protein has homology to galanin		
CC	receptor, chemokine receptor and somatostatin. GPCR 14273 sequences are		
CC	useful as a target for diagnosis and treatment of receptor-mediated		
CC	disorders. The polypeptide is used for screening of drugs, agonist,		
CC	antagonists and to produce antibodies. The antibodies are used for		
CC	diagnosis, to assess abnormal tissue distribution or expression during		
CC	development, inhibiting receptor function, tissue typing and forensic		
CC	identification. The DNA sequences are used to detect genetic alterations		
CC	in the gene and as a source of primers and probes		
XQ	Sequence 1561 BP; 321 A; 453 C; 386 G; 400 T; 0 U; 1 Other;		

Query Match	77.7%;	Score 1212;	DB 3;	Length 1561;	
Similarity	99.7%;	Pred. No. 0;			
Best Local	Matches 1462;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1	TTGGCAAGCTCAGCGTAAGCCTCTTTCACATGTCGAACTTCACAGAAAGGGGTTTCATGAGATGC	60		
Db	1	TTGGCAAGCTCAGCGTAAGCCTCTTTCACATGTCGAACTTCACAGAAAGGGGTTTCATGAGATGC	60		
QY	61	TTCAACACCATCAGTACGACCTCCAGACTTTGTCCGGCTTTTACCCGAACTTTTCAACGCGAG	120		
Db	61	TTCAACACCATCAGTACGACCTCCAGACTTTGTCCGGCTTTTACCCGAACTTTTCAACGCGAG	120		
QY	121	TCGATGACCCCTCTTGACAGACGACGAGGCGCGAGCTCCGCACTTCCCGGACGCGGAG	180		
Db	121	TCGATGACCCCTCTTGACAGACGACGAGGCGCGAGCTCCGCACTTCCCGGACGCGGAG	180		
QY	181	GCCAGGCGCCCGGAGATGTCCTCCATGATGTGACAGAGACGAGCGGCCTTGATCCCTCGACCA	240		
Db	181	GCCAGGCGCCCGGAGATGTCCTCCATGATGTGACAGAGAGAGGGGCTGATCCCTCGACCA	240		
QY	241	CCCTGACCAAATGCAATCGACACCACTTCCCTTTCTGTGGATGTCAAGGGCGACACAC	300		
Db	241	CCCTGACCAAAGTCAATCGACACCACTTCCCTTTCTGTGGATGTCAAGGGCGACACAC	300		
QY	301	GATTGTGTGTAAGGGTGTGTGAGACCAACCGTCTGAGGAACTCATCTTTGTGTCTCACTGC	360		
Db	301	GATTGTGTGTAAGGCTGTGTGAGACCAACCGTCTGAGGAACTCATCTTTGTGTCTCACTGC	360		
QY	361	TGGGCAACAGTGTGTCTCTAGTGTGTGTGGCGGCGCTCGGCGCGCGTGTGGGCGTCAAGCA	420		
Db	361	TGGGCAACAGTGTGTCTCTAGTGTGTGTGGCGGCGCGCTCGGCGCGCGTGTGGGCGTCAAGCA	420		
QY	421	GCTGTGTCTCAACCTCTTCTGCGCGGATTTGCTTTTCAACGAGCGCAATCCCTTAATGC	480		
Db	421	GCTGTGTCTCAACCTCTTCTGCGCGGATTTGCTTTTCAACGAGCGCAATCCCTTAATGC	480		
QY	481	TGCTGTGTGTGAGGCTGTGAGGCGCTGTGTGGGGCCCGTGTGTGCACTGCTCTTCT	540		
Db	481	TGCTGTGTGTGAGGCTGTGAGGCGCTGTGTGGGGCCCGTGTGTGCACTGCTCTTCT	540		
QY	541	ACGATGATGACATAGAGGCGGAGCGCTCAAGATCTTCACATGTGGCGCGGTCAAGCTGAGGC	600		
Db	541	ACGATGATGACATAGAGGCGGAGCGCTCAAGATCTTCACATGTGGCGCGGTCAAGCTGAGGC	600		
QY	601	GCATGTGTGTGATGTGTGCGCCTTCGGGCGCGGCTTTGAGCGGCGCGGCGGAGACTCAGG	660		
Db	601	GCATGTGTGTGATGTGTGCGCCTTCGGGCGCGGCTTTGAGCGGCGCGGCGGAGACTCAGG	660		
QY	661	CGGACATGTGTGCTTTTCAATATGAGGTTAATCTGTGGGCGCTTCGCGCGCTGCCCTTCAACT	720		
Db	661	CGGACATGTGTGCTTTTCAATATGAGGTTAATCTGTGGGCGCTTCGCGCGCTGCCCTTCAACT	720		
QY	721	TGTTCCGCGGAGTGTCCGAGAGGCTTCCCGGCGGAGACAGAAATTCGATTTTGGACAT	780		
Db	721	TGTTCCGCGGAGTGTCCGAGAGGCTTCCCGGCGGAGACAGAAATTCGATTTTGGACAT	780		
QY	781	TGATTTGGCCCAACCGCATAGAGAAATCTTCATGAGATGTGTTTTGTGACATTTGAAT	840		
Db	781	TGATTTGGCCCAACCGCATAGAGAAATCTTCATGAGATGTGTTTTGTGACATTTGAAT	840		
QY	841	TCTGTGTGTCCGGGACTGTGTATGTGATCACTTCAAAATTTTACATTCACGAAAG	900		
Db	841	TCTGTGTGTCCGGGACTGTGTATGTGATCACTTCAAAATTTTACATTCACGAAAG	900		
QY	901	CATGCGGAGAGAGGCTTACGCTGAGCTTTGGATATCTGTGAGGCGACAGATCCGAGTGT	960		
Db	901	CATGCGGAGAGAGGCTTACGCTGAGCTTTGGATATCTGTGAGGCGACAGATCCGAGTGT	960		
QY	961	CCCAACAAGACTACGAGCTCTTCCGAGAGGCTCTTCTGTCAATGATGTTTCTTTTCAATCA	1020		
Db	961	CCCAACAAGACTACGAGCTCTTCCGAGAGGCTCTTCTGTCAATGATGTTTCTTTTCAATCA	1020		
QY	1021	TGTGGAGTCCCATCATCAACATCTCTCTCATCTTGTATCAAAATCTTCCGAGAGAC	1080		

```

Db 1021 TGTGAGTCCCATCATCATCACTCCTCCTCATCTTGATCAAACTTCGGAGAGACC 1080
Qy 1081 TGTGATCTGAGCCATCCCTTTCTTCTGGGTGGTGGCTTCACTTTGCCAACTCTGCC 1140
Db 1081 TGTGATCTGAGCCATCCCTTTCTTCTGGGTGGTGGCTTCACTTTGCCAACTCTGCC 1140
Qy 1141 TAAACCCCATCTGTACAACTGTGCTGTTCAGAAAGAAATGAGAAATTTTGTCT 1200
Db 1141 TAAACCCCATCTGTACAACTGTGCTGTTCAGAAAGAAATGAGAAATTTTGTCT 1200
Qy 1201 GCTTCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 GCTTCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1261 TGTCTGTATTTTCCAGCTAACTAGCCTCTGTGTGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 TGTCTGTATTTTCCAGCTAACTAGCCTCTGTGTGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 GGGAGTTAACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GCGAGTTAACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1441 TCCTTAAAGAACTTCTATGAGTTCC 1467
Db 1441 TCCTTAAAGAACTTCTATGAGTTCC 1467

```

RESULT 4
ID AAA64344 standard; DNA; 180 BP.

AAA64344;

20-DEC-2000 (first entry)

Probe for murine G-protein coupled receptor designated cDNA.

Human; G-protein coupled receptor; receptor-mediated disorder;

14273 receptor; cardiovascular disease; congestive heart failure;

cardiac myocyte hypertrophy; probe; ss.

Homo sapiens.

WO200050596-A2.

31-AUG-2000.

28-FEB-2000; 2000MO-US005068.

26-FEB-1999; 99US-00261599.

08-DEC-1999; 99US-00456455.

(MILL-) MILLENIUM PHARM INC.

Glucksmann MA, Tsai F;

WPI; 2000-587184/55.

14273 receptor polynucleotides and polypeptides, useful in the diagnosis

and treatment of receptor-mediated disorders such as cardiovascular

diseases.

Example 2; Page 77; 105pp; English.

The present sequence represents a probe for cDNA encoding a murine G-

protein coupled receptor, designated 14273 receptor. The G-protein

coupled receptor 14273 polypeptide is used to produce antibodies, in drug

screening assays, and to screen for agonist and antagonists. The

antibodies are used for diagnostic applications, and to assess abnormal
tissue distributions or abnormal expression during development. The
antibodies are also useful for tissue typing, and in forensic
identification, as well as for inhibiting receptor function. The 14273
polynucleotides can be used to express the protein, to detect 14273 mRNA,
to detect genetic alterations in the 14273 gene, and as a source of
primers and probes. The 14273 polynucleotides and polypeptides are also
useful as a target for diagnosis and treatment of receptor-mediated
disorders, especially cardiovascular diseases such as congestive heart
failure caused by cardiac myocyte hypertrophy

Sequence 180 BP; 37 A; 54 C; 36 G; 53 T; 0 U; 0 Other;

Query Match 11.5%; Score 180; DB 3; Length 180;

Best Local Similarity 100.0%; Pred. No. 2e-63;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1021 TGTGAGTCCCATCATCATCACTCCTCCTCATCTTGATCAAACTTCGGAGAGACC 1080

Db 1 TGTGAGTCCCATCATCATCACTCCTCCTCATCTTGATCAAACTTCGGAGAGACC 60

Qy 1081 TGTGATCTGAGCCATCCCTTTCTTCTGGGTGGTGGCTTCACTTTGCCAACTCTGCC 1140

Db 61 TGTGATCTGAGCCATCCCTTTCTTCTGGGTGGTGGCTTCACTTTGCCAACTCTGCC 120

Qy 1141 TAAACCCCATCTGTACAACTGTGCTGTTCAGAAAGAAATGAGAAATTTTGTCT 1200

Db 121 TAAACCCCATCTGTACAACTGTGCTGTTCAGAAAGAAATGAGAAATTTTGTCT 180

RESULT 5
ID AAA64345 standard; DNA; 138 BP.

AAA64345;

20-DEC-2000 (first entry)

Probe for murine G-protein coupled receptor designated cDNA.

Human; G-protein coupled receptor; receptor-mediated disorder;

14273 receptor; cardiovascular disease; congestive heart failure;

cardiac myocyte hypertrophy; probe; ss.

Homo sapiens.

WO200050596-A2.

31-AUG-2000.

28-FEB-2000; 2000MO-US005068.

26-FEB-1999; 99US-00261599.

08-DEC-1999; 99US-00456455.

(MILL-) MILLENIUM PHARM INC.

Glucksmann MA, Tsai F;

WPI; 2000-587184/55.

14273 receptor polynucleotides and polypeptides, useful in the diagnosis

and treatment of receptor-mediated disorders such as cardiovascular

diseases.

Example 2; Page 77; 105pp; English.

The present sequence represents a probe for cDNA encoding a murine G-

protein coupled receptor, designated 14273 receptor. The G-protein

coupled receptor 14273 polypeptide is used to produce antibodies, in drug

screening assays, and to screen for agonist and antagonists. The

CC antibodies are also useful for tissue typing, and in forensic
CC identification, as well as for inhibiting receptor function. The 14273
CC polynucleotide can be used to express the protein, to detect 14273 mRNA,
CC to detect genetic alterations in the 14273 gene, and as a source of
CC primers and probes. The 14273 polynucleotide and polypeptides are also
CC useful as a target for diagnosis and treatment of receptor-mediated
CC disorders, especially cardiovascular diseases such as congestive heart
CC failure caused by cardiac myocyte hypertrophy
XX

Sequence 138 BP; 30 A; 46 C; 26 G; 36 T; 0 U; 0 Other;
SQ

Query Match 8.8%; Score 138; DB 3; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.8e-46;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 909 AAGAGGCTTACGCTGAGCTTGAGTCTGAGAGCCACGATCCGAGTGTCCCAACAA 968
DB 1 AAGAGGCTTACGCTGAGCTTGAGTCTGAGAGCCACGATCCGAGTGTCCCAACAA 60
QY 969 GACTACCGACTCTCCGACGCTCTTCTGCTCATGTTCTTTCATCATGTGAGT 1028
DB 61 GACTACCGACTCTCCGACGCTCTTCTGCTCATGTTCTTTCATCATGTGAGT 120
QY 1029 CCCATCATCATCACCATC 1046
DB 121 CCCATCATCATCACCATC 138
XX

RESULT 6
ABT04869 standard; cDNA; 1086 BP.
ID ABT04869 standard; cDNA; 1086 BP.
XX
AC ABT04869;
XX
DT 11-OCT-2002 (first entry)
XX
DE Human G protein coupled receptor hRUP31 coding sequence.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP33; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0255366P.
PR 20-FEB-2001; 2001US-0270266P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282356P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282365P.
PR 14-MAY-2001; 2001US-0280917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
PA (AREN-) ARENA PHARM INC.
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
XX WPI; 2002-566565/60.
DR P-PSDB; ABJ04071.
XX
XX Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
XX Claim 15; Page 61; 84pp; English.

CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR coding sequence of the invention
XX

Sequence 1086 BP; 185 A; 346 C; 295 G; 260 T; 0 U; 0 Other;
SQ

Query Match 2.8%; Score 44; DB 6; Length 1086;
Best Local Similarity 100.0%; Pred. No. 4.6e-08;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 852 GGACTGCTCATTTGATGATCACTTACTCCAAATTTTACAGATCAC 895
DB 658 GGACTGCTCATTTGATGATCACTTACTCCAAATTTTACAGATCAC 701
XX

RESULT 7
ABT73398 standard; DNA; 1086 BP.
ID ABT73398 standard; DNA; 1086 BP.
XX
AC ABT73398;
XX
DT 04-DEC-2002 (first entry)
XX
DE DNA encoding human GPCR HF1948 mutant I128F.
XX
KW Human; transmembrane receptor; G-protein coupled receptor; GPCR; allergy;
KW hypertension; reflux disease; depression; migraine; schizophrenia; ulcer;
KW psychiatric disorder; asthma; bronchospasm; anaesthesia;
KW myocardial infarction; MI; stroke; glaucoma; anxiety;
KW prostatic hyperplasia; epilepsy; prostate cancer; rhinitis; angina;
KW prostatic hypertrophy; mutant; ds.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200268600-A2.
XX
PD 06-SEP-2002.
XX
PF 26-FEB-2002; 2002WO-US005625.
XX
PR 26-FEB-2001; 2001US-0271913P.
XX
PA (AREN-) ARENA PHARM INC.
PI Liaw CW, Chalmers DT, Behan DP, Maciejewski-Lenior D, Leonard JN;
PI Lin I, Ortuno D;
XX
XX WPI; 2002-706980/76.
DR P-PSDB; ABG95169.
XX
PT New human G-protein coupled receptor (GPCR), useful for screening agonist
PT or inverse agonist compounds for treating diseases associated with GPCR.
XX
PS Example 2; Page 180-181; 201pp; English.
XX
XX The present invention relates to transmembrane receptors, particularly
CC endogenous human G-protein coupled receptors (GPCRs), mutant (non-
CC endogenous) versions of the GPCRs, and the polynucleotide sequences
CC encoding them. The GPCRs are useful for screening agonist or inverse
CC agonist compounds for treating diseases associated with GPCR. Diseases
CC that can be treated with such compounds include allergies, hypertension,
CC reflux disease, depression, migraine, schizophrenia, ulcer, psychotic
CC disorder, asthma, bronchospasm, anaesthesia, myocardial infarction (MI),
CC stroke, glaucoma, prostatic hyperplasia, epilepsy, prostate cancer,
CC anxiety, prostatic hypertrophy, rhinitis, and angina. The present
CC sequence encodes a mutant human GPCR
XX
SQ Sequence 1086 BP; 181 A; 348 C; 295 G; 262 T; 0 U; 0 Other;
Query Match 2.8%; Score 44; DB 6; Length 1086;

KM nototropic; neuroprotective; antiparkinsonian; antidiabetic;
 KM antidepressant; hypotensive; tranquilizer; nephrotropic;
 KM cell growth disorder; neurological condition; pituitary gland disorder;
 KM colon disorder; breast disorder; lung disorder; prostate disorder;
 KM Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;
 KM colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;
 KM sleeplessness; hypertension; anxiety; stress; renal failure; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1083
 FT /tag= a
 FT /product= "HGPRBM18"
 FT /note= "Human G protein coupled receptor"
 PN US2003022186-A1.
 XX 30-JAN-2003.
 PD 14-NOV-2001; 2001US-00992331.
 PF 14-NOV-2000; 2000US-0248483P.
 XX 16-JAN-2001; 2001US-0261782P.
 PR 27-JUL-2001; 2001US-0308540P.
 XX (FEDE/) FEDER J N.
 PA (MINT/) MINTIER G.
 PA (RAMA/) RAMANATHAN C S.
 PI Feder JN, Mintier G, Ramanathan CS;
 PT WPI; 2003-416985/39.
 DR P-PSDB; ABU09715.
 XX New isolated human GPCR, HGPRBM18, polynucleotide and polypeptide,
 PT useful for treating, diagnosing and preventing disorders such as cancer,
 PT neurological conditions, and diseases of the pituitary gland, colon,
 PT breast and lungs.
 XX Claim 1; Fig 1; 66pp; English.
 PS The invention describes a new isolated polynucleotide encoding a human G-
 CC protein coupled receptor (GPCR) or its functional fragment. The methods
 CC and compositions of the present invention are useful for treating,
 CC diagnosing, preventing and screening disorders associated with aberrant
 CC cell growth, neurological conditions, and diseases related to the
 CC pituitary gland, colon, breast, lungs and prostate. These disorders
 CC include Alzheimer's, Parkinson, diabetes, dwarfism, colour blindness,
 CC retinal pigmentosa, asthma, depression, schizophrenia, sleeplessness,
 CC hypertension, anxiety, stress and renal failure. This sequence encodes
 CC novel human G protein coupled receptor HGPRBM18
 XX
 SQ Sequence 1086 BP; 182 A; 349 C; 295 G; 260 T; 0 U; 0 Other;
 Query Match 2.8%; Score 44; DB 7; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 4.6e-08;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 852 GGAAGTGTGATTCAGTACTCCAAATTTTACAGATC 895
 DB 658 GGAAGTGTGATTCAGTACTCCAAATTTTACAGATC 701
 RESULT 11
 ADB47641 standard; cDNA; 1086 BP.
 AC ADB47641;
 XX 04-DEC-2003 (first entry)
 XX Human cDNA encoding GPCR, HGPRBM18.
 DE

XX Human; ss; gene; G protein-coupled receptor; GPCR; endocrine disorder;
 KM pituitary disorder; growth hormone; prolactin; luteinising hormone;
 KM follicle-stimulating hormone; thyroid-stimulating hormone;
 KM adrenocorticotropin; vasopressin; oxytocin; aberrant growth;
 KM aberrant lactation; aberrant sexual characteristic development;
 KM testosterone; oestrogen; aberrant water homeostasis; hypogonadism;
 KM Addison's disease; hypothyroidism; Cushing's disease; agromegaly;
 KM gigantism; lethargy; osteoporosis; aberrant calcium homeostasis;
 KM aberrant potassium homeostasis; reproductive disorder;
 KM developmental disorder; colon cancer; breast cancer; prostate cancer;
 KM lung cancer.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1086
 FT /tag= b
 FT /product= "HGPRBM18"
 FT /transl_except= (pos:364..372,aa:Gly-Val)
 FT misc_feature 1..1083
 FT /tag= a
 FT /note= "Claimed in claim 1"
 FT misc_feature 4..1083
 FT /tag= c
 FT /note= "Claimed in claim 1"
 PN US2003129653-A1.
 XX 10-JUL-2003.
 PD 30-SEP-2002; 2002US-00262313.
 PF 16-JAN-2001; 2001US-0261782P.
 XX 27-JUL-2001; 2001US-0308540P.
 PR 14-NOV-2001; 2001US-00992331.
 XX (FEDE/) FEDER J N.
 PA (MINT/) MINTIER G.
 PA (RAMA/) RAMANATHAN C S.
 PA (HAWK/) HAWKEN D R.
 PI Feder JN, Mintier G, Ramanathan CS, Hawken DR;
 PT WPI; 2003-678603/64.
 DR P-PSDB; ADB47642.
 XX New human G-protein coupled receptor, HGPRBM18, useful for treating and
 PT diagnosing disease such as cancer, hypothyroidism, Cushing's disease and
 PT osteoporosis.
 XX Claim 1; Fig 1; 71pp; English.
 PS The invention relate to an isolated human G protein-coupled receptor
 CC (GPCR), HGPRBM18, polynucleotide encoding a novel human GPCR HGPRBM18.
 CC Also included are expression vectors, host cells, a fusion protein
 CC comprising HGPRBM18 and an Fc portion of a human immunoglobulin protein,
 CC an anti-HGPRBM18 antibody its antigenic epitope, screening a library of
 CC molecules or compounds with a polynucleotide to identify at least one
 CC molecule or compound which specifically binds to the polynucleotide
 CC sequence and screening for candidate compounds capable of modulating
 CC activity of a G-protein coupled receptor-encoding polypeptide. HGPRBM18
 CC or its (ant)agonist (small molecule, peptide, and antisense molecule) is
 CC useful for treating a disease, disorder, or condition related to the
 CC endocrine, gastrointestinal, reproductive, pulmonary, or neural system.
 CC Diseases include endocrine disorders, disorders of the pituitary,
 CC aberrant growth hormone synthesis and/or secretion, aberrant prolactin
 CC synthesis and/or secretion, aberrant luteinising hormone synthesis and/or
 CC secretion, aberrant follicle-stimulating hormone synthesis and/or
 CC secretion, aberrant thyroid-stimulating hormone synthesis and/or
 CC secretion, aberrant adrenocorticotropin synthesis and/or secretion,
 CC aberrant vasopressin secretion, aberrant oxytocin secretion, aberrant
 CC growth, aberrant lactation, aberrant sexual characteristic development,

XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.

PS Disclosure; Fig 1; 523pp; English.

XX
CC The present invention describes antigenic peptides (1) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention

XX
SQ Sequence 1737 BP; 377 A; 487 C; 440 G; 430 T; 0 U; 3 Other;

Query Match 2.8%; Score 44; DB 7; Length 1737;

Best Local Similarity 100.0%; Pred. No. 4,4e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 852 GGACTGTCATTGTGATCACTTCTCCAAATTTTACAGATCAC 895
DB 701 GGACTGTCATTGTGATCACTTCTCCAAATTTTACAGATCAC 744

Search completed: September 30, 2004, 22:12:48

Job time : 572.119 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 21:52:37 : Search time 731.589 Seconds
(without alignments)
10810.058 Million cell updates/sec

Title: US-10-077-698-5

Perfect score: 1560
Sequence: 1 ttgcgaagcttcgagcgaagc.....aaaaaaaaaaaaaaaaaaaa 1560

Scoring table: OLIGO_NTIC
Gapop 60.0 , Gapext 60.0

Searched: 3340653 seqs, 2534783454 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1560	100.0	1560	US-10-086-181-4	Sequence 4, Appl1
2	1560	100.0	1560	US-10-077-698-5	Sequence 5, Appl1
3	1560	100.0	1560	US-10-171-027-5	Sequence 5, Appl1
4	1560	100.0	1560	US-10-075-987-5	Sequence 5, Appl1
5	1086	69.6	1086	US-10-086-181-6	Sequence 6, Appl1
6	181	11.6	181	US-10-171-027-14	Sequence 14, Appl1
7	138	8.8	138	US-10-171-027-15	Sequence 15, Appl1
8	57	3.7	259	US-09-535-459-1891	Sequence 1891, Ap
9	57	3.7	261	US-09-535-459-1881	Sequence 1881, Ap
10	57	3.7	285	US-09-535-459-1892	Sequence 1892, Ap
11	56	3.6	269	US-09-535-459-1882	Sequence 1882, Ap
12	50	3.2	262	US-09-535-459-1884	Sequence 1884, Ap
13	44	2.8	280	US-09-535-459-1889	Sequence 1889, Ap
14	44	2.8	1086	US-09-995-225-7	Sequence 7, Appl1

15	44	2.8	1086	10	US-09-992-331-1	Sequence 1, Appl1
16	44	2.8	1086	10	US-09-995-225-7	Sequence 7, Appl1
17	44	2.8	1086	14	US-10-086-181-3	Sequence 3, Appl1
18	44	2.8	1086	15	US-10-083-168-11	Sequence 11, Appl1
19	44	2.8	1086	15	US-10-083-168-78	Sequence 78, Appl1
20	44	2.8	1086	15	US-10-083-168-80	Sequence 80, Appl1
21	44	2.8	1086	15	US-10-262-313-1	Sequence 1, Appl1
22	44	2.8	1086	17	US-10-768-878-1	Sequence 1, Appl1
23	44	2.8	1160	14	US-10-015-498-1	Sequence 1, Appl1
24	44	2.8	1737	15	US-10-225-567A-681	Sequence 681, App
25	44	2.8	1743	14	US-10-086-181-1	Sequence 1, Appl1
26	44	2.8	1743	15	US-10-077-698-2	Sequence 2, Appl1
27	44	2.8	1743	15	US-10-171-027-2	Sequence 2, Appl1
28	44	2.8	1743	15	US-10-075-987-2	Sequence 2, Appl1
29	42	2.7	297	10	US-09-535-459-1895	Sequence 1895, Ap
30	40	2.6	1066	15	US-10-017-161-1809	Sequence 1809, Ap
31	40	2.6	1066	16	US-10-292-798-1465	Sequence 1465, Ap
32	32	2.6	1104	10	US-09-791-932-10	Sequence 10, Appl1
33	33	2.6	3173	15	US-10-116-252-5	Sequence 5, Appl1
34	39	2.5	638	9	US-09-828-644-58	Sequence 58, Appl1
35	38	2.4	203	10	US-09-535-459-1916	Sequence 1916, Ap
36	38	2.4	217	10	US-09-535-459-1915	Sequence 1915, Ap
37	38	2.4	221	10	US-09-535-459-1919	Sequence 1919, Ap
38	38	2.4	241	9	US-09-879-536-472	Sequence 472, App
39	38	2.4	250	10	US-09-535-459-1923	Sequence 1923, Ap
40	38	2.4	254	10	US-09-535-459-1921	Sequence 1921, Ap
41	38	2.4	273	10	US-09-535-459-1926	Sequence 1926, Ap
42	38	2.4	278	10	US-09-535-459-1913	Sequence 1913, Ap
43	38	2.4	280	10	US-09-535-459-1910	Sequence 1910, Ap
44	38	2.4	281	10	US-09-535-459-1920	Sequence 1920, Ap
45	38	2.4	283	10	US-09-535-459-1925	Sequence 1925, Ap

ALIGNMENTS

RESULT 1	US-10-086-181-4	Sequence 4, Application US/10086181
Publication No.	US20020177151A1	
GENERAL INFORMATION:		
APPLICANT:	GIMENO, Ruth	
TITLE OF INVENTION:	METHODS FOR THE TREATMENT OF METABOLIC DISORDERS, INCLUDING OBESITY AND DIABETES	
FILE REFERENCE:	NMI-220	
CURRENT APPLICATION NUMBER:	US/10/086,181	
CURRENT FILING DATE:	2002-02-26	
PRIOR APPLICATION NUMBER:	60/271,655	
PRIOR FILING DATE:	2001-02-26	
NUMBER OF SEQ ID NOS:	16	
SOFTWARE:	FASTSEQ for Windows Version 4.0	
SEQ ID NO 4		
LENGTH:	1560	
TYPE:	DNA	
ORGANISM:	Murine ortholog	
FEATURE:		
NAME/KEY:	CDS	
LOCATION:	(195)...(1280)	
US-10-086-181-4		
Query Match	100.0%;	Score 1560; DB 14; Length 1560;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 1560;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	TTGCAGAGCTCAGCGTAAGCTCTTCACATGCAATCTCAGAGGAGGTTTCATGAGTGC 60
DB	1	TTGCAGAGCTCAGCGTAAGCTCTTCACATGCAATCTCAGAGGAGGTTTCATGAGTGC 60
QY	61	TTGACACCATCAGTACGACCTCCAGACTTGTCCGGCTTTACCCGAATCTTCACAGCGGAG 120
DB	61	TTGACACCATCAGTACGACCTCCAGACTTGTCCGGCTTTACCCGAATCTTCACAGCGGAG 120
QY	121	TTGATGACCTCTTGAACAGCAGCAGAGCGGCGGAGCTCCGCATCTTCCGGAGCGGTGG 180

121 TCATGACCCCTTGAAGACGACGAGCGGAGCTCCGCACTCTCCGGAGCGGTGG 180
181 GCCGGGCGCCCGGAGATGCTCCCTGAGTGTGACAGAGAGCGGCGCTGTGCTCCGACA 240
181 GCCGGGCGCCCGGAGATGCTCCCTGAGTGTGACAGAGAGCGGCGCTGTGCTCCGACA 240
241 CCCGAGCAAGTGAATGAGACCCACTTCCCTTCTCTGAGATGTCAAGGGGAGACGACC 300
241 CCTGGACCAAGTGAATGAGACCCACTTCCCTTCTCTGAGATGTCAAGGGGAGACGACC 300
301 GGTGTGTGTGAGAGGTGAGAGACCAACGTTTCTGAGGACTCATTTGTGCTCACTGAC 360
301 GGTGTGTGTGAGAGGTGAGAGACCAACGTTTCTGAGGACTCATTTGTGCTCACTGAC 360
361 TGGGCAACGTGTGTGCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
361 TGGGCAACGTGTGTGCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
421 GCCGTGTGTCAACCTTCTGTGCGGAGATTGCTTCAACAGGAGCATCCCTGTAGTGC 480
421 GCCGTGTGTCAACCTTCTGTGCGGAGATTGCTTCTTCAACAGGAGCATCCCTGTAGTGC 480
481 TCGTGTGTGTGAGAGGCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
481 TCGTGTGTGTGAGAGGCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
541 ACCGTGTGTGAGAGGAGGAGGAGCTCAAGATCTTCACTGAGGAGGAGGAGGAGGAGG 600
541 ACCGTGTGTGAGAGGAGGAGGAGGAGCTCAAGATCTTCACTGAGGAGGAGGAGGAGG 600
601 GATGTGTGTGATGT 660
601 GATGTGTGTGATGT 660
661 GATGTGTGTGATGT 720
661 GATGTGTGTGATGT 720
721 TGT 780
721 TGT 780
781 TGT 840
781 TGT 840
841 TGT 900
841 TGT 900
901 CATGT 960
901 CATGT 960
961 CCCAACAAGACTACGACTCTTCCGCAAGCTCTTCTGCTGCTCATGTGTGTGTGTGT 1020
961 CCCAACAAGACTACGACTCTTCCGCAAGCTCTTCTGCTGCTCATGTGTGTGTGTGT 1020
1021 TGT 1080
1021 TGT 1080
1081 TGT 1140
1081 TGT 1140
1141 TAAACCCCATCTGTACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
1141 TAAACCCCATCTGTACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
1201 GCTTCTTTTTCAGAGAGGAGCAATTTTTCAGATACGTGTGTGTGTGTGTGTGTGTGT 1260
1201 GCTTCTTTTTCAGAGAGGAGCAATTTTTCAGATACGTGTGTGTGTGTGTGTGTGTGT 1260

1201 GCTTCTTTTTCAGAGAGGAGCAATTTTTCAGATACGTGTGTGTGTGTGTGTGTGTGT 1260
1261 TGT 1320
1261 TGT 1320
1321 GGGAGTTAACTTCAAGGAAAGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
1321 GGGAGTTAACTTCAAGGAAAGCCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
1381 ACAGAGGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
1381 ACAGAGGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
1441 TCCCTTAAAGAACTTCTATAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
1441 TCCCTTAAAGAACTTCTATAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
1501 GATCTAGTTAATAATTTTATTTATATATATATATATATATATATATATATATATAT 1560
1501 GATCTAGTTAATAATTTTATTTATATATATATATATATATATATATATATATATAT 1560

RESULT 2
US-10-077-698-5
; Sequence 5, Application US/10077698
; Publication No. US20030008350A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14273 Receptor, A No. US20030008350A1 G-Protein Coupled Receptor
; FILE REFERENCE: 5800-4B, 035800/177086
; CURRENT APPLICATION NUMBER: US/10/077,698
; PRIORITY FILING DATE: 2002-02-13
; PRIORITY APPLICATION NUMBER: 09/261,599
; PRIORITY FILING DATE: 1999-02-26
; PRIORITY APPLICATION NUMBER: 09/107,761
; PRIORITY FILING DATE: 1998-06-30
; PRIORITY APPLICATION NUMBER: 09/223,538
; PRIORITY FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Murine ortholog
US-10-077-698-5

Query Match 100.0%; Score 1560; DB 15; Length 1560;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TTGCAAGCTCACGCTTACGCTTTCACATGCAATCTCAGAGAGGGGTTTATGAGATGC 60
1 TTGCAAGCTCACGCTTACGCTTTCACATGCAATCTCAGAGAGGGGTTTATGAGATGC 60
61 TTCAACCATCAAGTACCACTTCCGAGCTTTCACGAGATCTTTCACAGCGGAG 120
61 TTCAACCATCAAGTACCACTTCCGAGCTTTCACGAGATCTTTCACAGCGGAG 120
121 TCATGACCCCTTGAAGACGACGAGCGGAGCTCCGCACTCTCCGGAGCGGTGG 180
121 TCATGACCCCTTGAAGACGACGAGCGGAGCTCCGCACTCTCCGGAGCGGTGG 180
181 GCCGGGCGCCCGGAGATGCTCCCTGAGTGTGACAGAGAGCGGCGCTGTGCTCCGACA 240
181 GCCGGGCGCCCGGAGATGCTCCCTGAGTGTGACAGAGAGCGGCGCTGTGCTCCGACA 240
241 CCCGAGCAAGTGAATGAGACCCACTTCCCTTCTCTGAGATGTCAAGGGGAGACGACC 300
241 CCCGAGCAAGTGAATGAGACCCACTTCCCTTCTCTGAGATGTCAAGGGGAGACGACC 300
301 GGTGTGTGTGAGAGGTGAGAGACCAACGTTTCTGAGGACTCATTTGTGCTCACTGAC 360
301 GGTGTGTGTGAGAGGTGAGAGACCAACGTTTCTGAGGACTCATTTGTGCTCACTGAC 360

QY	541	ACGTAATGACAAATGAGCCGACGCGTCAAGATCCCTCACTAGCCGCGGTGACCTCGAAGC	600
Dp	541	ACGTAATGACAAATGAGCCGACGCGTCAAGATCCCTCACTAGCCGCGGTGACCTCGAAGC	600
QY	601	GCAATGATGTGATCGTGTGCGCTCCGCGCGCGCTTGAGAGGCGCCCGAGGCGAGACTAGG	660
Dp	601	GCAATGATGTGATCGTGTGCGCTCCGCGCGCGCTTGAGAGGCGCCCGAGGCGAGACTAGG	660
QY	661	CGGCACTGTGAGCTTTCATATGAGGGTTACTGTGGCGCTCGCGCGCTGCCCTTACATCT	720
Dp	661	CGGCACTGTGAGCTTTCATATGAGGGTTACTGTGGCGCTCGCGCGCTGCCCTTACATCT	720
QY	721	TGTTCCGCGTGTGCCCGAGCGCCCTCCCGCGGGGAGCAGAGAAATTCGATTTGGACAT	780
Dp	721	TGTTCCGCGTGTGCCCGAGCGCCCTCCCGCGGGGAGCAGAGAAATTCGATTTGGACAT	780
QY	781	TGATATGGCCCCAACCGCATAGAGAAATCTATGAGATGTGTGTTTGGACTTTGAACT	840
Dp	781	TGATATGGCCCCAACCGCATAGAGAAATCTATGAGATGTGTGTTTGGACTTTGAACT	840
QY	841	TCTGTGTGTCGGGAGACTGTCTATTGTGATCACTTACTCCAAAATTTTACAGATCAGAAAG	900
Dp	841	TCTGTGTGTCGGGAGACTGTCTATTGTGATCACTTACTCCAAAATTTTACAGATCAGAAAG	900
QY	901	CATGCGGAGAAAGGCTTACGCTGAGCTGTGGATATCTCGAAGGCCACAGATCCGAGGT	960
Dp	901	CATGCGGAGAAAGGCTTACGCTGAGCTGTGGATATCTCGAAGGCCACAGATCCGAGGT	960
QY	961	CCCAAACAAGACTACGACTCTTTCGGCAGCGCTCTTCTGTGCTAGGTTCCTTCTTCATCA	1020
Dp	961	CCCAAACAAGACTACGACTCTTTCGGCAGCGCTCTTCTGTGCTAGGTTCCTTCTTCATCA	1020
QY	1021	TGTGAGATGCCATATCATATCAACATCCCTGCTCATCTTGAATCCAAAACCTTCCGGAGAGCC	1080
Dp	1021	TGTGAGATGCCATATCATATCAACATCCCTGCTCATCTTGAATCCAAAACCTTCCGGAGAGCC	1080
QY	1081	TGATCATGTGCGCACATCCCTTTTCTCTGTGGGTGTGGGCGCTTCAGCTTTGGCCAACTCGGCC	1140
Dp	1081	TGATCATGTGCGCACATCCCTTTTCTCTGTGGGTGTGGGCGCTTCAGCTTTGGCCAACTCGGCC	1140
QY	1141	TAAACCCCATATCTGTACCAACATGTGCTGTTTCAGGAAGAAATGAGAAATTTTGTCT	1200
Dp	1141	TAAACCCCATATCTGTACCAACATGTGCTGTTTCAGGAAGAAATGAGAAATTTTGTCT	1200
QY	1201	GCTTCTTTTTCACAGAGAAAGGAGCATTTTTCAGATAGCTGTCTCAGCGGAAATGACT	1260
Dp	1201	GCTTCTTTTTCACAGAGAAAGGAGCATTTTTCAGATAGCTGTCTCAGCGGAAATGACT	1260
QY	1261	TGTCTGTATTTTCCAGCTACTAGCTCTGTGTGCGAGGTGAAACAACGCTGTGACATGTAA	1320
Dp	1261	TGTCTGTATTTTCCAGCTACTAGCTCTGTGTGCGAGGTGAAACAACGCTGTGACATGTAA	1320
QY	1321	GGGAGTTACTTTCAGAGAAAGCCACACAGTGGCTCTGCTTTTAAATAATACCGCACTTCCA	1380
Dp	1321	GGGAGTTACTTTCAGAGAAAGCCACACAGTGGCTCTGCTTTTAAATAATACCGCACTTCCA	1380
QY	1381	ACAGACGAGCATCTACGAGCCAGCAAAATTAAGAAATGATGCTCAGTATTAATAAATATTTT	1440
Dp	1381	ACAGACGAGCATCTACGAGCCAGCAAAATTAAGAAATGATGCTCAGTATTAATAAATATTTT	1440
QY	1441	TCCTTAAAGAACTTTATGAGGTTCCTTTTGTGAACCTTTTAAAGTGTGTTGTATAT	1500
Dp	1441	TCCTTAAAGAACTTTATGAGGTTCCTTTTGTGAACCTTTTAAAGTGTGTTGTATAT	1500
QY	1501	GATCAGTAAATTAATTTTATTTTATTTATACAGTGTCTCTACAAAAAATTTTTTTTTTTT	1560
Dp	1501	GATCAGTAAATTAATTTTATTTTATTTATTTATACAGTGTCTCTACAAAAAATTTTTTTTTT	1560

RESULT 4
US-10-075-987-5
; Sequence 5, Application US/10075987
; Publication No. US20030166061A1

```

; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14273 Receptor, A No. US20030166061A1el G-Protein Coupled Receptor
; FILE REFERENCE: 5800-4B, 035800/177086
; CURRENT APPLICATION NUMBER: US/10/075,987
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US/09/261,599B
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/223,538
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Murine ortholog
US-10-075-987-5

Query Match      100.0%; Score 1560; DB 15; Length 1560;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

	Query Match	Similarity	Score	DB	Length	Best Local	Mismatches	Indels	Gaps
		100.0%	1560	15	1560	Local Similarity	Conservative		
		100.0%	Pred. No.			Matches	1560		
OY	1	TTCGCAACCTCAGCTTAAGCCCTCTTCCACTGCATCTCACAGAAAGGGATTTCATGAGATGC	60						
Db	1	TTGCCAAGCTCAGCGTTAAGCCCTCTTCCACTGCATCTCACAGAAAGGGATTTCATGAGATGC	60						
OY									
Db	61	TTTCACACATCATGTGACCACTTCACAGACTTTGTCCGCTTTACCCTGAATCTTTCACAGCGAG	120						
OY									
Db	61	TTTCACACATCATGTGACCACTTCACAGACTTTGTCCGCTTTACCCTGAATCTTTCACAGCGAG	120						
OY	121	TTCGATGACCCCTCTTTCACAGACACAGAGCGCGGCGACCTCCGACATCTTCCCGAGCGGTGG	180						
Db	121	TTCGATGACCCCTCTTTCACAGACACAGAGCGCGGCGACCTCCGACATCTTCCCGAGCGGTGG	180						
OY	181	GCCGGGCGCCCCGGGCAATGTCCCCTGAGTGTGACAAGACGAGGGGCCCTGTGTCTCTTGACA	240						
Db	181	GCCGGGCGCCCCGGGCAATGTCCCCTGAGTGTGACAAGACGAGGGGCCCTGTGTCTCTTGACA	240						
OY	241	CCCTGAGACCAAGTCAAATGGCACCCACTTCCCTTTCTTCTCGGATGTCCMAAGGCGGACAC	300						
Db	241	CCCTGAGACCAAGTCAAATGGCACCCACTTCCCTTTCTTCTCGGATGTCCMAAGGCGGACAC	300						
OY	301	GGTTGGTGTMGAGCSTCGGAGACACACGTTCTGAGACTCAATCTTGTGCTGTCACTGC	360						
Db	301	GGTTGGTGTMGAGCSTCGGAGACACACGTTCTGAGACTCAATCTTGTGCTGTCACTGC	360						
OY	361	TGGGCAACGTGTGTCTCTAAGTCTGTGTGGCGGCGGCTGTGGCGGCGGTGGGGCGTACCCA	420						
Db	361	TGGGCAACGTGTGTCTCTAAGTCTGTGTGGCGGCGGCTGTGGCGGCGGTGGGGCGTACCCA	420						
OY	421	GGCTGGTCTCAACTCTTTCGGCGGGGANTTGTCTTCAACAGCGGCAATCCCTCTAAGTGC	480						
Db	421	GGCTGGTCTCAACTCTTTCGGCGGGGANTTGTCTTCAACAGCGGCAATCCCTCTAAGTGC	480						
OY	481	TGATGTGTGCTGAGACTGAGGCGTGTGGGCGCCGTCGTCTGCAACTGTCTTCT	540						
Db	481	TGATGTGTGCTGAGACTGAGGCGTGTGGGCGCCGTCGTCTGCAACTGTCTTCT	540						
OY	541	ACGTATATGACAATGAGCGGCGACGCTACAGATCTCTCACTATGGCGCGGCTCAAGCTGGAAC	600						
Db	541	ACGTATATGACAATGAGCGGCGACGCTACAGATCTCTCACTATGGCGCGGCTCAAGCTGGAAC	600						
OY	601	GCATGTGTGCAATCGTAGGCGCTCCGAGCGGCTTGAAGCGGCCCGGAGCGGCGGACTCAGG	660						
Db	601	GCATGTGTGCAATCGTAGGCGCTCCGAGCGGCTTGAAGCGGCCCGGAGCGGCGGACTCAGG	660						
OY	661	CGGCACTGTGCTTTCAATATGGGGTTACTGGCGCTCGCGCGCTGCCCTCTACATCT	720						
Db	661	CGGCACTGTGCTTTCAATATGGGGTTACTGGCGCTCGCGCGCTGCCCTCTACATCT	720						
OY	721	TGTTCCGGTGTGTCCTCGGACGCGCTTCCGAGCGGGGACACAGAAATTCGATTTGCAAT	780						
Db	721	TGTTCCGGTGTGTCCTCGGACGCGCTTCCGAGCGGGGACACAGAAATTCGATTTGCAAT	780						

Qy 1215 GAGAGGAGGACATTTTACAGATAGCTGTGAGGGAAATGACTTGTCTATTTC 1274
 |||
 Db 1021 GAGAGGAGGACATTTTACAGATAGCTGTGAGGGAAATGACTTGTCTATTTC 1080

Qy 1275 AGCTAA 1280
 |||
 Db 1081 AGCTAA 1086

RESULT 6

US-10-171-027-14
 ; Sequence 14, Application US/10171027
 ; Publication No. US20030073168A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria A.
 ; APPLICANT: Tsai, Fong-Ying
 ; TITLE OF INVENTION: 14273 Receptor, A No. US20030073168A1 G-Protein Coupled Recepto
 ; FILE REFERENCE: NMI-204CP3
 ; CURRENT APPLICATION NUMBER: US/10/171,027
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US/09/456,455
 ; PRIOR FILING DATE: 1999-12-08
 ; PRIOR APPLICATION NUMBER: 09/223,538
 ; PRIOR FILING DATE: 1998-12-30
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 181
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-10-171-027-14

Query Match 11.6%; Score 181; DB 15; Length 181;
 Best Local Similarity 100.0%; Pred. No. 1.5e-82;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1021 TGTGAGTCCCATCATCATCCATCCCTCCATCTTGATCCAAACTTCGCGAGGACC 1080
 |||
 Db 1 TGTGAGTCCCATCATCATCCATCCCTCCATCTTGATCCAAACTTCGCGAGGACC 60

Qy 1081 TGTGATCTGGCCATCCCTTTCTTCTGGGTGTGGCTTCACGTTTGCACACTCTGCC 1140
 |||
 Db 61 TGTGATCTGGCCATCCCTTTCTTCTGGGTGTGGCTTCACGTTTGCACACTCTGCC 120

Qy 1141 TAAACCCCATCTGTACAACTGTGCTGTTCAGGAAGCATGAGGAAGATTTTGGCT 1200
 |||
 Db 121 TAAACCCCATCTGTACAACTGTGCTGTTCAGGAAGCATGAGGAAGATTTTGGCT 180

Qy 1201 G 1201
 |
 Db 181 G 181

RESULT 7

US-10-171-027-15
 ; Sequence 15, Application US/10171027
 ; Publication No. US20030073168A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria A.
 ; APPLICANT: Tsai, Fong-Ying
 ; TITLE OF INVENTION: 14273 Receptor, A No. US20030073168A1 G-Protein Coupled Recepto
 ; FILE REFERENCE: NMI-204CP3
 ; CURRENT APPLICATION NUMBER: US/10/171,027
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US/09/456,455
 ; PRIOR FILING DATE: 1999-12-08
 ; PRIOR APPLICATION NUMBER: 09/223,538
 ; PRIOR FILING DATE: 1998-12-30
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 138

; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-10-171-027-15

Query Match 8.8%; Score 138; DB 15; Length 138;
 Best Local Similarity 100.0%; Pred. No. 2.3e-60;
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 909 AAGAGCTTACGCTTGAGCTTGAGCACTCTGAGAGCCAGATCCGAGTCCCAACAA 968
 |||
 Db 1 AAGAGCTTACGCTTGAGCTTGAGCACTCTGAGAGCCAGATCCGAGTCCCAACAA 60

Qy 969 GACTACCGACTCTTCGCGAGCTCTTCCTCATGTGTTCTTCTTCATCATGTGAGT 1028
 |||
 Db 61 GACTACCGACTCTTCGCGAGCTCTTCCTCATGTGTTCTTCTTCATCATGTGAGT 120

Qy 1029 CCCATCATCATCAACCATC 1046
 |||
 Db 121 CCCATCATCATCAACCATC 138

RESULT 8

US-09-535-459-1891/C
 ; Sequence 1891, Application US/09535459
 ; Publication No. US20030040615A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sellhammer, Jeffrey J.
 ; APPLICANT: Delegeane, Angelo M.
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Stuve, Laura L.
 ; APPLICANT: Naughton, Rebecca E.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULES
 ; FILE REFERENCE: PD-1014 CIP
 ; CURRENT APPLICATION NUMBER: US/09/535,459
 ; CURRENT FILING DATE: 2000-03-24
 ; Prior application data removed - consult PAM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2170
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1891
 ; LENGTH: 259
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030040615A1 rat00161335
 ; LOCATION: 104, 238
 ; OTHER INFORMATION: a, c, g, or other
 US-09-535-459-1891

Query Match 3.7%; Score 57; DB 10; Length 259;
 Best Local Similarity 100.0%; Pred. No. 1.6e-18;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 AGCTCTTCACATGTCATCTCACAGAGGGGTTTCATGAGTGTTCACACCATCAGT 74
 |||
 Db 59 AGCTCTTCACATGTCATCTCACAGAGGGGTTTCATGAGTGTTCACACCATCAGT 3

RESULT 9

US-09-535-459-1881
 ; Sequence 1881, Application US/09535459
 ; Publication No. US20030040615A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sellhammer, Jeffrey J.
 ; APPLICANT: Delegeane, Angelo M.
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Stuve, Laura L.
 ; APPLICANT: Naughton, Rebecca E.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULES
 ; FILE REFERENCE: PD-1014 CIP

```

; CURRENT APPLICATION NUMBER: US/09/535,459
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2170
; SOFTWARE: PERL Program
; SEQ ID NO 1881
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030040615A1 rat00212947
; NAME/KEY: unsure
; LOCATION: 23, 81, 171
; OTHER INFORMATION: a, t, c, g, or other
US-09-535-459-1881

Query Match      3.7%; Score 57; DB 10; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.6e-18;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GCCTCTTCACATGCAATCTCAGAGAGGGGTTTCATGAGTCTTCACACCATCAGTG 75
DB 24 GCCTCTTCACATGCAATCTCAGAGAGGGGTTTCATGAGTCTTCACACCATCAGTG 80

RESULT 10
US-09-535-459-1892/c
; Sequence 1892, Application US/09535459
; Publication No. US20030040615A1
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULE
; FILE REFERENCE: PD-1014 CIP
; CURRENT APPLICATION NUMBER: US/09/535,459
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2170
; SOFTWARE: PERL Program
; SEQ ID NO 1892
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030040615A1 rat00190334
; NAME/KEY: unsure
; LOCATION: 158, 160, 237, 267
; OTHER INFORMATION: a, t, c, g, or other
US-09-535-459-1892

Query Match      3.7%; Score 57; DB 10; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.6e-18;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 AGCCTCTTCACATGCAATCTCAGAGAGGGGTTTCATGAGTCTTCACACCATCAGT 74
DB 107 AGCCTCTTCACATGCAATCTCAGAGAGGGGTTTCATGAGTCTTCACACCATCAGT 51

RESULT 11
US-09-535-459-1882
; Sequence 1882, Application US/09535459
; Publication No. US20030040615A1
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
US-09-535-459-1882
```

```

; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULE
; FILE REFERENCE: PD-1014 CIP
; CURRENT APPLICATION NUMBER: US/09/535,459
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2170
; SOFTWARE: PERL Program
; SEQ ID NO 1882
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030040615A1 rat00209119
; NAME/KEY: unsure
; LOCATION: 23, 25, 126
; OTHER INFORMATION: a, t, c, g, or other
US-09-535-459-1882

Query Match      3.6%; Score 56; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 5.2e-18;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CTCTTCACATGCAATCTCAGAGAGGGGTTTCATGAGTCTTCACACCATCAGTGA 76
DB 26 CTCTTCACATGCAATCTCAGAGAGGGGTTTCATGAGTCTTCACACCATCAGTGA 81

RESULT 12
US-09-535-459-1884/c
; Sequence 1884, Application US/09535459
; Publication No. US20030040615A1
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULE
; FILE REFERENCE: PD-1014 CIP
; CURRENT APPLICATION NUMBER: US/09/535,459
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2170
; SOFTWARE: PERL Program
; SEQ ID NO 1884
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030040615A1 rat00083975
; NAME/KEY: unsure
; LOCATION: 2
; OTHER INFORMATION: a, t, c, g, or other
US-09-535-459-1884

Query Match      3.2%; Score 50; DB 10; Length 262;
Best Local Similarity 100.0%; Pred. No. 6.5e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 AGTCGATGACCCCTTCAGACGACGAGCGGCGGACCTCGGCATCTTC 168
DB 110 AGTCGATGACCCCTTCAGACGACGAGCGGCGGACCTCGGCATCTTC 61

RESULT 13
US-09-535-459-1889/c
; Sequence 1889, Application US/09535459
```

```
/ Publication No. US20030040615A1
/ GENERAL INFORMATION:
/ APPLICANT: Sellhammer, Jeffrey J.
/ APPLICANT: Delegeane, Angelo M.
/ APPLICANT: Stuart, Susan G.
/ APPLICANT: Stuve, Laura L.
/ APPLICANT: Mullahy, Sara J.
/ APPLICANT: Naughton, Rebecca E.
/ TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULE
/ FILE REFERENCE: PD-1014 CIP
/ CURRENT APPLICATION NUMBER: US/09/535,455
/ CURRENT FILING DATE: 2000-03-24
/ Prior application data removed - consult PAM or file wrapper
/ NUMBER OF SEQ ID NOS: 2170
/ SOFTWARE: PERL Program
/ SEQ ID NO 1889
/ LENGTH: 280
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Inctye ID No. US20030040615A1 rat00204309
/ NAME/KEY: unsure
/ LOCATION: 2, 20, 230, 260, 268, 273
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-535-459-1889
```

```
Query Match 2.8%; Score 44; DB 10; Length 280;
Best Local Similarity 100.0%; Pred. No. 8.2e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 31 GCATCTCAGAGAGGGTTCATGAGTGTTCACACCATCAGT 74
Db 144 GCATCTCAGAGAGGGTTCATGAGTGTTCACACCATCAGT 101
```

```
RESULT 14
US-09-995-225-7
/ Sequence 7, Application US/09995225
/ Publication No. US20020193584A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Kuoping
/ APPLICANT: Chu, Zhi Liang
/ APPLICANT: Dang, Huong T.
/ APPLICANT: Lowitz, Kevin P.
/ APPLICANT: Pride, Geron
/ TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human C
/ TITLE OF INVENTION: Receptors
/ FILE REFERENCE: AREN-0308
/ CURRENT APPLICATION NUMBER: US/09/995, 225
/ CURRENT FILING DATE: 2001-11-26
/ PRIOR APPLICATION NUMBER: 09/170,496
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: PCT/US99/23938
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: 60/253,404
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/255,366
/ PRIOR FILING DATE: 2000-12-12
/ PRIOR APPLICATION NUMBER: 60/270,286
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,365
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/270,266
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,032
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,358
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,356
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
```

```
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 7
/ LENGTH: 1086
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20020193584A1e1 Sequence
/ US-09-995-225-7
```

```
Query Match 2.8%; Score 44; DB 9; Length 1086;
Best Local Similarity 100.0%; Pred. No. 8.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 852 GCACTGTCATTGTGATCAGTTACTCCAAATTTTACAGATCAG 895
Db 658 GCACTGTCATTGTGATCAGTTACTCCAAATTTTACAGATCAG 701
```

```
RESULT 15
US-09-992-331-1
/ Sequence 1, Application US/09992331
/ Publication No. US20030022186A1
/ GENERAL INFORMATION:
/ APPLICANT: FEDER, JOHN N.
/ APPLICANT: MINTIER, GABE
/ APPLICANT: RAMANATHAN, CHANDRA S.
/ APPLICANT: HAMKEN, DONALD R.
/ TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HPRERNY18,
/ TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
/ FILE REFERENCE: D0048NP
/ CURRENT APPLICATION NUMBER: US/09/992,331
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/308,540
/ PRIOR FILING DATE: 2001-07-27
/ PRIOR APPLICATION NUMBER: 60/261,782
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: 60/248,483
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 45
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 1086
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-992-331-1
```

```
Query Match 2.8%; Score 44; DB 10; Length 1086;
Best Local Similarity 100.0%; Pred. No. 8.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 852 GCACTGTCATTGTGATCAGTTACTCCAAATTTTACAGATCAG 895
Db 658 GCACTGTCATTGTGATCAGTTACTCCAAATTTTACAGATCAG 701
```

```
Search completed: October 1, 2004, 04:22:21
Job time : 733.589 secs
```